Title: Moving Beyond End-to-End Path Information to Optimize CDN Performance

Authors: Rupa Krishnan, Harsha V. Madhyastha, Sridhar Srinivasan, Sushant Jain, Arvind Krishnamurthy, Thomas Anderson, Jie Gao

Venue: ACM IMC 2009 pp. 190-201

Reviewers: Jack Kosaian and Allison McDonald

I. Summary

In this paper, the authors challenge the idea that redirecting clients to Content Distribution Network (CDN) nodes with the lowest latency is sufficient for producing optimal client interactions. Measuring differences between minimum round-trip times (RTTs) among members of an address prefixes showed that a significant number of client paths had inflated latencies despite being connected to the geographically closest CDN node. Through their measurement study, the authors attribute this inflated latency to circuitous routes and queueing delay. To assist network administrators in discovering and mitigating the causes of latency inflation, the authors propose *WhyHigh*. *WhyHigh* uses a combination of passive data analysis and active route probing to determine sets of prefixes that are undergoing significant latency inflation. The system can rank the severity of latency inflation and propose the underlying causes of circuitous routes, as well as means for network administrators to mitigate the issues. In evaluating *WhyHigh*, the system was used by Google to mitigate several production latency inflation issues due to a variety of causes.

II. Strengths

1. Combination of measurement study with system/solution design

The authors challenge the assumption that a CDN choosing the node with the lowest latency will produce the best result for a client by conducting a measurement study of the various factors at play in client latency experience. While this measurement study in itself provided valuable insight into the limitations of current CDN node selection, the authors also spent significant effort on identifying the causes of the observed latency problems and building a tool that could help the network administrators at Google further identify and address the causes of the latency. This deviates from the traditional outline followed by many measurement studies which often perform a measurement, analyze results, and list implications without offering a systematic solution for mitigating any issues found. The authors provide a much more holistic study in developing a diagnostic tool based on their measurements.

2. *Evaluation in a production setting*

The authors deployed *WhyHigh* for use by network administrators at Google to determine the tool's effectiveness in assisting in the mitigation of observed latency inflation. This represents a major strength of the paper for multiple reasons. Many systems and solutions proposed in academia are evaluated in a synthetic manner either through simulation or controlled replication of real-world observations. This generally makes it challenging to view the benefits of using particular systems in production settings. By employing *WhyHigh* for administrators facing live issues, the true utility of *WhyHigh* as a diagnostic assistant can be better determined. Further, given the scale at which Google operates, one would imagine that the company has rather advanced administration tools. The fact that *WhyHigh* was determined to aid administrators operating at such scale helps to improve one's opinion about the tool's utility.

3. *Combination of active and passive measurement methods*

In the design of *WhyHigh*, the authors utilized both passive measurement tools such as BGP paths and active measurement tools like traceroute and ping. While this might increase the difficulty and complexity of *WhyHigh*, using both active and passive measurement techniques gives a more dynamic and complete pictures of the network topology, giving the authors more power to diagnose why certain clients experience unreasonably high latencies.

III. Improvements and Extensions

Despite the strengths of the paper presented above, there are a number of ways in which this work could be improved upon for greater clarity or extended in future studies:

1. *A more diverse sample*

The authors describe the measurement dataset in section 2.4. The entirety of the dataset was collected on two days which were three months apart. Given large-scale implications of latency inflation and the cost and challenge of correcting the problem, an additional longitudinal study may have shed light on the changes in latency inflation over time. In fact, the paper entirely lacks a discussion on the length of time latency inflation must occur to make the problem consequential, how common latency inflation is for short durations, and how much fluctuation in latency occurs from day-to-day or hour-to-hour. The paper also did not discuss how latency varies throughout a single day, for example if RTTs increase during business hours, which the authors might have been able to speculate about considering their conclusion that queueing delay was a significant contributor to inflated latencies.

Additionally, of the 173K prefixes they identify, the authors proceed to eliminate 82K of them for not fitting their sampling model — for example, having inconclusive geolocation data, not being served by the nearest CDN, or representing too large a geographic area. A further discussion would be welcome of how these excluded data points could be used to help measure network latencies generally and how they might have been used in *WhyHigh*, even if the inclusion would have been prohibitive in terms of complexity and the prefixes were to remain excluded.

2. *The role of queueing delay*

In section 3.3, the authors discuss the role of queueing in the observed latency inflation. After measuring the changes in routes across two successive days and observing no significant increase in latency inflation for those prefixes whose routes have changed, they conclude that a large portion of the inflated latency they observe is due to queueing. Queueing delay is certainly difficult to measure, and identifying where in the route the queueing is occurring is equally challenging. Even if queueing were easier to identify, it is difficult to mitigate considering the large number of independent parties involved and the nature of the problem (e.g. hardware limitations). But while the authors claim queueing delay to be the largest contributor to latency inflation, they then proceed to ignore the issue of queueing delay entirely in their *WhyHigh* design. The analysis of their tool would have been stronger if the authors had discussed possible solutions to large queueing delays or how queueing delay might fit into an administrator's role of diagnosing and remedying inflated latencies.

3. *Further evaluation of* WhyHigh*'s effectiveness*

Because the authors took on the task of building a tool to identify the root cause of latency inflation in the CDN *as they occur*, evaluating the effectiveness of the tool is one of the most difficult tasks in the paper. Their current evaluation of *WhyHigh*, in which several discrete problems were identified and referred to an administrator, only shows anecdotal evidence of the usefulness of *WhyHigh* in reducing the instances of latency inflation in Google's CDN. A possible extension to this evaluation methodology could be utilizing Google's documentation on the previous methods for solving unusual latency inflation (a discussion which was lacking in the paper) to simulate or replicate scenarios in which *WhyHigh* should be able to identify the problem. This is a common methodology for evaluating new systems, although it certainly poses significant challenges when applied to large-scale networking systems.

4. *Effectiveness of* WhyHigh *as a tool for system administrators*

Though the authors' deployment of *WhyHigh* for real-world use at Google is commendable, the evaluation presented in the paper sheds little light on how the tool truly benefits network administrators. *WhyHigh*'s utility as a system administration tool would be greatly strengthened if the authors could provide even a rough estimate of the number of admin-hours saved by using *WhyHigh*. Papers on configuration management and performance issue mitigation often make very clear how challenging root cause analysis can be for even expert system administrators and developers. The authors of this paper do not provide such a view in their presentation of *WhyHigh*. Are the limited number of root causes of latency inflation pinpointed by *WhyHigh* things that network administrators at Google would have substantial difficulty identifying without the tool? If not, *WhyHigh*'s utility as a diagnostic tool is diminished.

5. *Predicting the side-effects of suggested actions*

After analyzing latency inflation at the prefix and AS level, and ranking measured entities by their degree of latency inflation, *WhyHigh* suggests the root cause of latency inflation for specified prefixes. The authors show that the root causes identified by *WhyHigh* generally have concrete means of mitigation. However, *WhyHigh*'s analysis stops at root cause identification. It would be interesting to extend *WhyHigh* to be able to project both the intended outcomes of mitigating a specific root cause as well as the unintended side-effects that could take place. For example, if *WhyHigh*'s analysis determines that changes in peering would reduce latency inflation, a more advanced tool might also simulate the effects of these policy changes on other communications. Perhaps a peering change that reduces latency inflation for one AS causes traffic to be routed circuitously from other ASes. Additionally, if *WhyHigh* suggests increasing bandwidth on a certain link to mitigate inflation latency, it would be useful to model changes in queueing delay that might occur due to this change. Given that queueing delay plays a substantial role in latency inflation, such an analysis would help to determine whether making a certain change would prove performant. Supporting these types of prediction would require *WhyHigh* to perform simulations based on the topology of interest and network conditions from historical logs.

6. Understanding the young-node dilemma

The authors note multiple times during the paper that prefixes served by CDN nodes that have been deployed for less than a year tend to have substantial latency inflation. The cause of this phenomenon is left unexplored by this paper, but presents an interesting grounds for future work. For example, one might perform a long-term study of the deployment of a new CDN node and the latencies of prefixes which it servers. Perhaps the inflated latencies experienced by these prefixes is due to misconfiguration, or perhaps it takes a significant amount of time for prefixes routes from prefixes to "converge" to use this node.