Controlling bandwidth in the clouds

August 29 2007

Looking to the Clouds, and finding answers. Barath Raghavan, a computer science Ph.D. candidate at UCSD's Jacobs School of Engineering looked to computing clouds and, together with is co-authors, developed a new way to manage bandwidth. Credit: UC San Diego

If half your company's bandwidth is allocated to your mirror in New York, and it's the middle of the night there, and your sites in London and Tokyo are slammed, that New York bandwidth is going to waste. UC San Diego computer scientists have designed, implemented, and evaluated a new bandwidth management system for cloud-based applications capable of solving this problem.

The UCSD algorithm enables distributed rate limiters to work together to enforce global bandwidth rate limits, and dynamically shift bandwidth allocations across multiple sites or networks, according to current network demand.

"With our system, an organization with mirrored Web sites or other services across the globe could dynamically shift its bandwidth allocations between sites based on demand. You can't do that now, and this lack of control is a significant drawback to today's cloud-based computing approaches," said Barath Raghavan, the first author on a new paper describing the work, and a Ph.D. candidate in the Department of Computer Science and Engineering at UCSD's Jacobs School of Engineering.

On August 30, this work will be presented in Kyoto, Japan at ACM SIGCOMM, one of the world's most prestigious networking conferences. It earned the 2007 SIGCOMM best student paper award – the top prize at the conference.

Raghavan's paper "Cloud Control with Distributed Rate Limiting," is co-authored with Kashi Vishwanath, Sriram Ramabhadran – two fellow UCSD computer science Ph.D. students – Kenneth Yocum (a UCSD computer science researcher) and Alex C. Snoeren, a computer science professor from UCSD's Jacobs School of Engineering.

The "flow proportional share" algorithm the UCSD computer scientists created enables the coordinated policing of a cloud-based service's network traffic, and therefore, the cost associated with this traffic. The TCP-centric design is scalable to hundreds of nodes, runs with very low overhead, and is robust to both loss and communication delay, making it practical for deployment in nationwide service providers, the authors write. TCP (Transmission Control Protocol) is the Internet protocol that establishes a connection between two hosts and ensures that packets travel safely from sender to receiver; TCP is used for the vast majority of Internet traffic.

“Our primary insight is that we can use TCP itself to estimate bandwidth demand,” said Alex Snoeren, the senior author on the paper. “Relying on TCP, we can provide the fairness that you would see with one central rate limiter. From the perspective of the actual network flows that are going through it, we have made our distributed rate limiter look centralized. This is the main technical contribution of the paper,” said Snoeren.

Distributed rate limiting could be useful in a variety of ways:

-- Cloud-based resource providers could control the use of network bandwidth, and associated costs, as if all the bandwidth were sourced from a single pipe.
-- For content-distribution networks that currently provide replication services to third-party Web sites, distributed rate limiting could provide a powerful tool for managing access to client content.
-- Internet testbeds such as Planetlab are often overrun with network demands from users. Distributed rate limiting could bring the bandwidth crisis under control and render such research tools much more effective.

Going with the (TCP) Flow

When you connect to a Web site, you open a TCP flow between your personal computing device and the Web site. As more people connect to the same site, the number of flows increases and the bandwidth gets split up fairly among the growing number of users. The flow proportional share algorithm from UCSD monitors the flows at each of the distributed sites and finds the largest flow at each site. Based on how fast packets are moving along the largest flow, the algorithm reverse engineers the network demand at each rate limiter.

A gossip protocol communicates these demand values among all the associated rate limiters. Next, the new algorithm ranks the rate limiters according to network traffic demands and splits the bandwidth accordingly.

“Our algorithm allows individual flows to compete dynamically for bandwidth not only with flows traversing the same limiter, but with flows traversing other limiters as well,” Raghavan explained.

The UCSD computer scientists performed extensive evaluations of their system, using Jain’s fairness index as a metric of inter-flow fairness. In a wide variety of Internet traffic patterns, the authors demonstrated that their approach to distributed limiting of the rate of bandwidth usage works as well as a single rate limiter would work, with very little overhead.

“As computing shifts to a cloud-based model, there is an opportunity to reconsider how network resources are allocated,” said Stefan Savage, a UCSD computer scientist not involved in the study.

Right now, Raghavan and co-author Kenneth Yocum are working on the code necessary to implement their flow proportional share algorithm within Planetlab, the Internet testbed that computer scientists use to test projects in a global environment.

Source: University of California - San Diego

Citation: Controlling bandwidth in the clouds (2007, August 29) retrieved 6 May 2024 from https://phys.org/news/2007-08-bandwidth-clouds.html