Research Statement

Mario A. Sánchez
Department of Electrical Engineering and Computer Science
Northwestern University
msanchez@eecs.northwestern.edu

My research interests focus on experimental systems deployment, performance analysis, network protocols and distributed Internet services such as DNS and content distribution networks (CDNs). I am interested in the application and implementation of low-level systems to analyze the performance of underlying protocols and their properties. I am currently working on two large research projects that entail evaluating the performance of layered network services from the perspective of individual peers in P2P systems located at the edge of the network.

DNS Performance and its impact on CDN services: The Domain Name Systems originally envisioned purpose of delegating the task of Internet name resolution among different distributed systems has expanded dramatically. Many new services have emerged that leverage the existing ubiquitous deployment of the DNS hierarchy to offer other functionality; one such service is offered by Content Distribution Network (CDN) systems. CDNs attempt to reduce latency in the delivery of content (improving overall performance) by relying on DNS name resolution to evaluate the clients proximity to the closest content replica server that will serve the request.

The emergence of such services has stressed the importance of DNS robustness and has redefined the term DNS reliability as well as the impact of configuration errors. The impact on a clients perceived quality-of-service from a timed-out request directed to a primary DNS server is no longer limited to an increased resolution delay while the request is redirected to a secondary DNS server (possibly located in a different network or even country). It also translates into an increased overall delay as the client is directed to a content replica sever that is as far away from it as the responding DNS server.

Previous DNS studies focus on lookup measurements, success rate, consistency and overall response time but due to the inherent difficulty of accessing a representative set of end nodes, analyzing the system from the edge of the network has proven to be a challenge. Taking advantage of a deployment of measurement software located on thousands of peers in a large-scale P2P network we are performing an extensive analysis on the reliability of the DNS hierarchy, evaluating the proximity of clients to their configured DNS servers as well as analyzing the effect of DNS configuration errors to ultimately weight-in the impact on these overlay networks.

Network Positioning at The Edge of the Network: Many existing network services such as VoIP, video streaming and file sharing systems rely on the capacity of clients to identify and contact the optimal (lowest-delay) node(s) from those capable of offering the required service. One way of finding such a node would entail directly probing each peer/server to identify the best match; however, this brute-force approach cannot scale to hundreds of thousands or millions of nodes in today's P2P systems. Many existing network positioning systems have emerged that attempt to estimate the network proximity of two nodes in the Internet without incurring on costly and bandwidth consuming direct latency measurements between them. The main limitation of these systems (studies) has been the evaluation of their performance in large-scale deployments. In this work, we explore the accuracy and overhead of existing network positioning systems based on the largest empirical dataset collected to date that incorporates measurements from the edge of the network, and find that their performance is much worse at scale than in simulation or limited testbed deployment. We are currently developing a new approach to network positioning that incorporates topological information to improve accuracy while retaining the appealing decentralized properties of previous network positioning systems.