My research career has focused on developing tools and techniques for troubleshooting the Internet. As an undergraduate I worked on measuring and diagnosing problems in the wide-area Internet, under the supervision of Professors Tom Anderson and Arvind Krishnamurthy at the University of Washington. At UC Berkeley I am working on making individual networks more robust. Two principles have guided me throughout all of these endeavors: I concentrate my efforts in areas where my work will have meaningful impact on the network community and the world at large, and I constantly seek out new experiences and skills that will help me become a more effective researcher.

Background and motivation: The commercial Internet was not designed with measurement and fault-diagnosis in mind; it provides limited primitives for troubleshooting problems, and leaves network operators with little ability to avoid or repair faults outside of their own network. My research seeks to address this problem with solutions that are deployable in the Internet today.

Reverse traceroute: traceroute is the first tool network operators use when a problem occurs, as it provides visibility into the path packets traverse to arrive at a given destination. While this information is highly valuable, it only provides half of the diagnostic picture, since most Internet communication is two-way, and the paths generally differ in the two directions. The first research project I participated in, known as reverse traceroute, aimed to complete that picture by revealing the return path used by arbitrary destinations.

My experiences with reverse traceroute have shaped my thinking in many ways. When I started on the project, I struggled to write simple programs. By the end I had obtained the skills necessary to understand and build large-scale distributed systems, perform measurement experiments, and work effectively as a team member. During this time I was able to make several research contributions, including a novel application of the set cover algorithm to network probing, which improved reverse traceorute's measurement completion time substantially, thereby bringing reverse traceroute from a proof-of-concept to a practical system. My contributions to reverse traceroute helped lead to a best paper in NSDI 2010 [1], and a real-world deployment actively used by network operators to measure Internet paths. For my contributions I also received honorable mention for the CRA outstanding undergraduate researcher award, and a Gary Kildall scholarship.

CloudFront: I applied my knowledge of Internet measurement as an intern at CloudFront, Amazon Web Service's content distribution network (CDN). CDNs exist to shorten the distance between users and content, thereby minimizing latency and decreasing traffic over expensive long-distance network links. Leveraging my familiarity with Internet behavior, I began by developing tools to aggregate, analyze, and trigger alarms for anomalous latency data, which served to help operators troubleshoot performance problems with their network. My internship culminated in a system to measure, prioritize and suggest fixes for *circuitous, suboptimal routes* between CloudFront and the thousands of client networks it serves. With this system, my team members could spend their time wisely on the routing issues that were most detrimental.

My team members at CloudFront taught me two immensely valuable skills. First and foremost, they showed me the importance of *guiding your work* based on feedback from the people who are ultimately going to benefit from it. Second, they taught me the importance of *collaborating with others* in order to generate ideas and identify important problems. During my internship I also had

the opportunity to give my first talk, after realizing that the CloudFront team would benefit from learning and using the reverse traceroute system I had helped build.

Wide Area Fault Isolation: During my internship I was stunned by how often my team had to fix connectivity problems with client networks. This lead me to ask: how prevalent are Internet outages, and what can be done about it? I started by measuring ping losses between Amazon's EC2 and a number of large Internet service providers. I found that long-lasting outages are surprisingly common; the data revealed more than 12,000 outages over two months More interestingly, I found that 80% of the outages were partial, where at least one vantage point maintained connectivity with the destination, suggesting that alternate, unused routes existed in the network.

With this as motivation, I began developing techniques to isolate the faults which are the cause of such outages, to help guide operators to fix them. Traditional tools such as traceroute provide insufficient information for this purpose, since they are blind to the *direction* of failure. By leveraging novel measurement tools such as spoofed forward traceroute, a historical path atlas, reverse path information, and a new algorithm to analyze these data, I was able to show that wide-area network faults can be isolated much more accurately than traditional tools allow. This work was awarded best senior thesis award for the class of 2011 [2]. It also formed a crucial component in our recent HotNets 2011 publication [3], and NSDI 2012 submission [4].

More than any of my previous experiences, my thesis work taught me to be an independent researcher. Under the mentorship of my advisors, I took a lead on this project, from formulating a research question, developing experiments, generating solutions, to communicating my results.

Current and Future Research: The main insight I took away from my previous research is that today's Internet is incredibly complex, and thus fundamentally difficult to manage and troubleshoot. At UC Berkeley I am now exploring ways to make the Internet more robust from *within* the network, by leveraging an emerging technology known as software-defined networking (SDN), a novel network architecture. This has given me an in-depth understanding of the potential benefits and shortcomings of SDN, and ultimately helped lead me to the realization that it could enable local entrepreneurs in the developing world, as described in the attached research proposal.

Broader impact: A common theme throughout my research has been an explicit decision to work on problems that are well-motivated and important to the communities around me. Reverse path information is highly valuable to network operators -- one company recently adopted reverse traceroute for their own operational use. By making it easier to debug performance and reliability problems, reverse traceroute consequently benefits the millions of users served by the Internet. More recently, I have begun thinking about how the difficulty of network management affects operators in developing regions. I am moving to address these problems with the goal of enabling the spread of Internet connectivity to under-privileged populations across the world.

References

[1] E. Katz-Bassett, H. Madhyastha, V. Adhikari, **C. Scott,** et. al. *Reverse Traceroute*, NSDI 2010

[2] **C. Scott,** *LIFEGUARD: Locating Internet Failure Events and Generating Usable Alternate Routes Dynamically*, Senior Thesis, 2011, http://www.eecs.berkeley.edu/~rcs/writeup.pdf

[3] E. Katz-Bassett, D. Choffnes, **C. Scott**, T. Anderson, A. Krishnamurthy, *Machiavellian Routing*, Hotnets 2011

[4] E.Katz-Basset, **C.Scott**, et. al, *Improving Internet Availability with LIFEGUARD*, *in submission*.