

CNS Core: Collaborative Research: Network Design Automation

When we use Office365, Amazon, or Facebook, we access a world of rich services via a network. Network reliability is paramount for business, science, and government because downtime affects productivity, economic output, and social communication. However, reliability is difficult to achieve in the face of component failure, manual configuration, and the stringent economics of the IT industry.

This grant seeks to take first steps towards creating a new field, *Network Design Automation* (NDA). NDA seeks to create computer-aided design (CAD) tools that reduce design effort and increase reliability based on a scientific understanding of network structure. We are inspired by the Electronic Design Automation (EDA), a \$7B industry that underpins the \$100B chip industry and is also a vibrant intellectual discipline in its own right. We will build on the last five years of progress in network verification done by the community that has already resulted in industry startups. However, existing work in network verification and synthesis primarily focuses on *router configuration files in data center networks*.

We seek to broaden the agenda to *debugging networked systems* and to other specialized networks such as *rural networks* and CDNs. Specific new tasks we will pursue include tying application performance to network problems using machine learning, understanding failure logs using Natural Language Processing, creating more controllable routers inspired by hardware boundary scan ideas such as JTAG, rural network design using constraint satisfaction, network scripting to lower the barriers for operators, and data mining to find configuration bugs without a specification. Just as the EDA industry has a suite of tools such as place-and-route, synthesis, and timing verification, the goal of NDA is to produce a *set of interconnected tools* to construct and manage networks that satisfy formal properties, while minimizing the frequency of failures and reducing recovery time.

UCLA will serve as the host institution. UCLA (Varghese: Networking; Millstein: Programming Languages, Netravali, Tamir: Systems) will be responsible for tools and theory for systems debugging, specification mining, and testing. USC (Govindan: Networking, Ragahavan: Computing for Social Good) will be responsible for network synthesis tools for other specialized networks such as rural networks. However, as we are all in Los Angeles we will all work on all problems together with regular meetings, co-advising students between the two institutions to expose synergies between research topics. Varghese and Govindan will coordinate individual research teams into a coherent whole, Netravali will lead the cross-cutting theme of systems debugging, and Ragahavan will lead the cross-cutting theme of rural network synthesis.

Intellectual Merit. Networks have several unique challenges that preclude reusing existing approaches in debugging and synthesis of programs and chips. These include scale (10^5 devices and zettabytes of traffic), heterogeneity (devices, protocols, vendors), management complexity (most configuration is still at "assembly" level), rate of evolution (new data centers and service rollouts), and high availability needs (<40 minutes per month downtime for 3×9 s availability). Thus, we need better abstractions and tools across the board: languages for modeling topologies, faults, and configurations, and tools for integrated debugging, specification mining, and configuration scripting. We seek to harness interdisciplinary approaches. Besides a consistent focus on using algorithm search (e.g., SAT solvers for rural topologies), we will leverage data mining (to learn specifications), machine learning (to build associations for debugging), NLP (to mine operator logs), and hardware test (to suggest new router primitives). We seek to not merely reuse existing mechanisms from these fields (e.g., SAT solvers, K-means) but to exploit domain-specific insights about the *structure* of modern networks to invent new scalable approaches such as *network specific data mining*.

Broader Impacts. We seek to transform academic, industrial and rural approaches to networking as follows: (1) Public Clouds: We will partner with Microsoft (via Albert Greenberg) and Google (via Amin Vahdat), former colleagues of Varghese and Govindan to develop, deploy, and evaluate NDA technology for clouds. (2) Rural Networks: Ragahavan will work rural networks (Motech in California, and Grameen in Ghana) to help deploy topology design and configuration scripting tools (3) Outreach to Operators: We will train underrepresented minorities by developing tools and workshops.

Keywords: [Compilers and Programming Languages, Distributed Systems] [Data Center Networks, Access Networks] [Manageability, Reliability, Robustness] [Algorithmic Search, Machine Learning, NLP]