

Research Statement

Boon Thau Loo, University of Pennsylvania

My research focuses on the development of new programming tools and analysis techniques that improve the process of designing, implementing, and securing large-scale distributed systems. A unifying theme is use of the novel *declarative networking* [CACM'09] framework, in which distributed systems are specified and implemented using a declarative recursive query language. Declarative networking fundamentally changes the way networking protocols are designed and implemented. My early work in declarative networking [SIGCOMM'05, SOSP'05] demonstrated its use as a domain specific language for fast prototyping, by compiling protocols written in a declarative language into efficient distributed implementations.

My work in the past three years has evolved declarative networking's original roots as a rapid prototyping framework, towards one that serves as an important bridge connecting formal theories (for reasoning about protocol correctness) and actual implementations. The ability to bridge this gap is a major step forward compared to traditional approaches, where formal specifications, proof of protocol correctness and implementations are decoupled from one another. This decoupling leads to increased development time, error-prone implementations, and tedious debugging. To test this new approach, I address four main challenges in different stages of the distributed systems development cycle:

- **Generating safe implementations.** The *FSR (Formally Safe Routing)* toolkit [HotNets'09, SIGCOMM'11 demo] is motivated by the complexity of policy configuration over the Internet's global routing system, whose convergence depends on how individual networks configure their Border Gateway Protocol (BGP) policies. FSR brings research in routing algebras into the context of declarative networking, to produce provably-correct distributed implementations. Our work on FSR has also resulted in formalizing the operational semantics of declarative networking programs [PPDP'11], and new distributed recursive view maintenance techniques [ICDE'09a best paper].
- **Declarative debugging of distributed systems.** *NetTrails* [SIGMOD'10b, SIGMOD'11 demo] is a declarative engine that allows a network operator to issue queries explaining the derivation and change of network state at any given node in a distributed system. NetTrails uses declarative networking as a basis for maintaining and querying network provenance at Internet-scale. By integrating NetTrails with tamper-evident logging, we recently showed that network provenance can be securely stored and queried even in the presence of Byzantine faults [SOSP'11].
- **Securing distributed systems.** The *DS2 (Declarative Secure Distributed Systems)* platform provides high-level programming abstractions for implementing secure distributed systems, achieved by unifying declarative networking and logic-based access control specifications [ICDE'09b]. DS2 has a wide range of applications, including reconfigurable trust management [CIDR'09], secure distributed data processing [SIGMOD'10a], and tunable anonymity [NDSS'10]. We are extending FSR's capabilities for analyzing secure protocols written in DS2, and performing safety/security analysis in the presence of incomplete routing information [WRiPE'11].
- **Optimizing distributed systems.** *PUMA (Policy-based Unified Multi-radio Architecture)* is a declarative constraint solving platform for optimizing wireless mesh protocols. Given high-level optimization goals and constraints as input, PUMA allows network operators to vary the choice of routing protocols [ICNP'09] and wireless channel selection [PRESTO'10]. We have also started some preliminary work on applying PUMA's declarative optimization platform to perform automated cloud resource orchestration [SOCC'11].

Impact. Declarative networking's widespread impact is shown not only in its publications across several subfields in computer science, but also its impact on industry and the research community at large. We have participated in multi-PI research initiatives that use declarative networking, such as DARPA's SAFER and WNaN programs on anonymous and adaptive wireless communications, and NEBULA, a NSF-funded multi-university effort led by Penn on a future Internet architecture for secure cloud computing. Overall, NSF has recognized the significance and inter-disciplinary nature of my work, by awarding nine grants (six as PI) related to declarative networking in the past 4.5 years.

We have also collaborated with several industry partners, including AT&T Research, Intel, LogicBlox, Microsoft Research, Raytheon BBN Technologies, and SRI International. Our work on DS2 [CIDR'09, SIGMOD'10a] has been integrated into LogicBlox's commercial data management system. I co-chaired two inter-disciplinary workshops (NetDB'09 and the inaugural WRiPE'11), co-located with SOSP and ICNP, which respectively explore the application of database and formal methods techniques in the networking domain.

NetDB@Penn. I lead the NetDB@Penn (<http://netdb.cis.upenn.edu>) research group, which comprises six doctoral students (three female), one post-doctoral researcher, and several masters students and undergraduates. I have graduated two doctoral students (one won the best thesis award in CS at Penn). I actively promote undergraduate research participation—undergraduate alumni from my group include a CRA honorable mention awardee, a Microsoft scholarship winner, and doctoral students in CS. An important aspect of my research is its integration into the curriculum of courses that I teach: I have used declarative networking in a large undergraduate/graduate networking course at Penn [SIGCOMM'11 Education]. I have started Ph.D. seminar courses at Penn that explore topics at the intersection of databases, networking, and formal methods.