Datacenter network latencies are approaching their microsecond-scale speed-of-light limit, and network bandwidths continue to grow beyond 100 Gbps. These improvements bear rethinking the design of communication-intensive distributed systems for datacenters, whose performance has historically been limited by slow networks. With the slowing down of Moore's law, a popular approach is to redesign distributed systems to use custom network hardware devices and technologies---smart network cards (NICs), lossless networks, programmable NICs, and programmable switches---that offload communication or data access from commodity CPUs.

In this dissertation, we show that we can continue to use end-to-end communication mechanisms to build high-performance distributed systems with commodity hardware in modern datacenters, i.e., we bring the speed of fast networks to distributed systems without requiring an expensive redesign with custom hardware. We show that the ubiquitous Remote Procedure Call (RPC) communication mechanism, when rearchitected specially for the capabilities of modern commodity datacenter hardware, is a fast, scalable, flexible, and simple communication choice for distributed systems. We make three contributions. First, we present a detailed analysis of datacenter communication hardware---ranging from the peripheral bus that connects CPUs to NICs, to the datacenter's switched network---that informs our choice of the communication mechanism. Second, we lay out the advantages of RPCs over in-network offloads through the design and evaluation of two new systems, a key-value store called HERD, and a distributed transaction processing system called FaSST. Third, we combine the lessons learned from the first two steps with new insights about datacenter packet loss and congestion control to create a new RPC library called eRPC, and show how existing distributed system codebases perform well over eRPC. In many cases, these systems substantially outperform offloads because they use less communication, and their end-to-end design provides flexibility and simplicity.

Thesis Committee:
David G. Andersen, Chair
Michael Kaminsky
Justine Sherry
Miguel Castro, Microsoft Research, Cambridge