

Special University Ph.D. Oral Examination
Load-Balancing and Parallelism for the Internet
Sundar Iyer

Department of Computer Science
Stanford University

10:00-11:00AM, Tuesday, February 18th, 2003
CISX Bldg, Room 101X (Cypress Auditorium)
(Refreshments served from 9:30am)

— (The talk is open to everyone) —

In this talk, I will describe how load-balancing and parallelism can be used in the design of networking systems that are scalable, fault tolerant and give performance guarantees. Today, due to the rapid growth of the Internet, networking systems are designed in a rather ad hoc manner. They are not easily scalable and provide little or no guarantees. As an example, most Internet routers do not give 100% throughput, i.e. they cannot achieve the capacity they were designed to achieve. The lack of scalability affects future growth, while the lack of fault tolerance and performance guarantees hampers the deployment of new applications such as telephony, streaming multimedia and broadcast on the Internet.

The Internet infrastructure consists of routers interconnected by links. In order to be scalable, fault tolerant and give performance guarantees, it is necessary that each of its components have the above capabilities. The components of concern to a packet on the Internet, can be broadly classified into the following areas: (1) Router Architecture, which deals with how to design the internals of a router, (2) Packet Processing, which deals with what arithmetic and storage operations need to be performed on a packet, (3) Packet Switching, which deals with how to efficiently move packets within a router and (4) Internet Routing, which deals with which link to transfer a packet over the Internet towards its final destination.

In my talk, I shall consider design issues in each of the above areas and present solutions by using load-balancing algorithms and readily available commodity parallel components. I will present data gathered from the network to motivate these solutions, show via theoretical analysis that they give performance guarantees, and touch upon their scalability and optimality. When required, I will also present simulation results to suggest their effectiveness. As a result, I will attempt to convince you that load-balancing and parallelism offer a practical way of supporting the future design requirements of the Internet, and will play a key role in its future growth and evolution.