Chadi BARAKAT
Chargé de recherche
INRIA – Planète research group
2004, route des Lucioles
06902 Sophia Antipolis - France

Email : Chadi.Barakat@sophia.inria.fr
Web : http://www.inria.fr/planete/chadi

Title of the course: Internet measurement and traffic analysis

Summary : The Internet has become a network of huge size that no person is able to control. Operators and end users are more and more in need for techniques to measure the traffic in the Internet, to infer its topology, to evaluate the performance of applications, and to detect any anomaly. The domain of Internet Measurement and Traffic Analysis has gained lot of importance the last years and most of the research labs in the world have activities in this direction. In this course, we summarize the different techniques used to monitor the Internet as well as the main tools. We mainly focus on passive and active measurement approaches and we give examples of each one. Then, we present an analysis of Internet traffic based on real measurements together with an overview of the models for Internet traffic. The course contains a refresh on the main protocols of the Internet and on the theoretical and statistical tools required to analyze the data collected by measurements.

Table of contents

    - SNMP
    - NetFlow
    - A list of important tools

- Measurement infrastructures
    - Introduction
    - Some Examples:
        - The Skitter infrastructure
        - The Route Views infrastructure
        - The Surveyor infrastructure
        - The Network Analysis Infrastructure
        - Internet Weather / Traffic Reports
        - The National Internet Measurement Infrastructure (NIMI)
- Inferring the Internet from end-to-end measurements
    - MINC: Multicast Inference of Network Characteristics
    - Neighborhood inference using delay measurements
    - Inference of Bottleneck characteristics from Poisson probes
    - Estimation of number of receivers in a multicast session from periodic probabilistic ACKs

- Inferring Internet topology using traceroute and BGP tables

Part II: Traffic analysis

- Introduction
    - Motivations for traffic analysis: Network dimensioning, modeling and simulation
    - Refresh on traffic measurement approaches
    - Summary of traces to be used (except for TCP models)

- Theoretical background
    - Common probability distributions
    - Distribution fitting
    - Heavy (or power) tails
    - Accuracy vs. Stationarity

- Contributions of different layers to Internet traffic
    - Traffic components
    - Impact of TCP congestion control
    - Traffic aggregation into flows

- Application behavior characteristics
    - Email
    - FTP
    - WWW access
    - Other applications: multimedia, games, peer-to-peer, etc.

- Self-similarity
    - Origins and definition
    - Long-range dependence

    - Estimation of the Hurst parameter
    - Construction of self-similar traffic
    - Consequences

- Dial-up user behavior

- Backbone measurements

- Traffic models
    - Introduction
    - Factors to account for when modeling traffic at different layers
    - Multilevel models: Examples
    - Modeling traffic at the flow level
    - TCP models
        - Introduction
        - A simple model for TCP throughput: Square Root Formula
        - Enhancing the square root formula to account for the packet nature of TCP, timeouts, and receiver window.
        - Advanced modeling of TCP.
        - Inferring the parameters for a model for TCP.
        - Modeling TCP using processor sharing queues