S-38.310 Thesis Seminar

Author:

Timo Viipuri

Title:

Internet Protocol Traffic Analysis for Network Simulation Purposes

Supervisor:

Jorma Jormakka (Prof.)

Instructor:

Marko Luoma (Lic. Tech.),
Networking laboratory, HUT

Context

- Background
- Objectives
- Network structure
- Measured parameters
- Traffic model
- Initial findings
- Conclusions
- Future work

Background

- Traffic is routinely analysed in nearly all operational networks
  - Monitoring user behaviour
  - Detecting faults
  - Analysing network performance
- Simulations are used to analyse network performance
  - Can be analysed before network is built
  - Middle ground between theory and practice
  - Accuracy and reliability of the results?

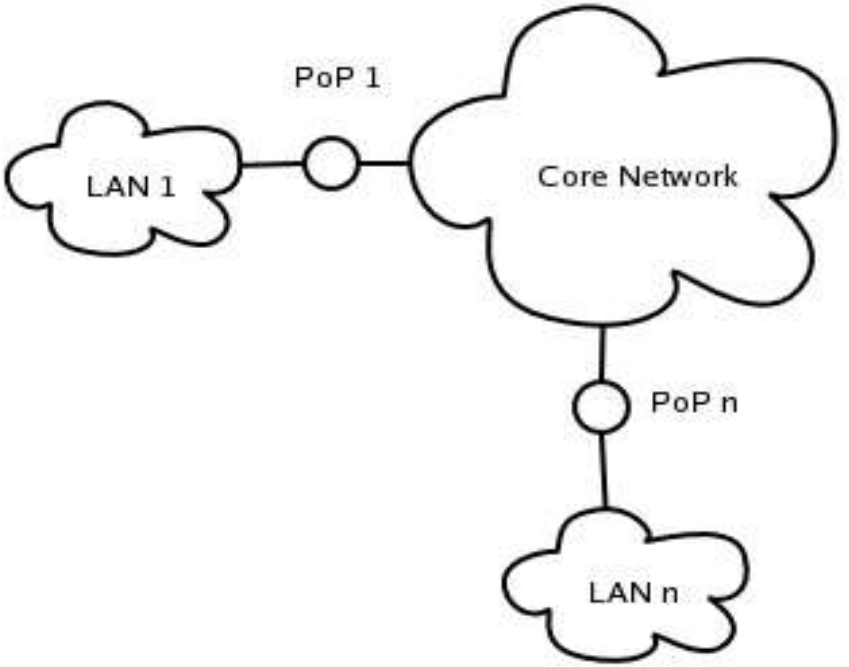
Objectives

- Analysing the network traffic
  - Traffic patterns
  - Network performance
  - Cost-effective
- Creating a traffic model
  - Used for replicating the traffic as accurately as possible in the simulation environment
  - Both qualitative and quantitative
  - Based on the traffic patterns

(Simulations are not part of this thesis)

Network Structure

- Core network with multiple PoPs (Point of Presence) to connect to LANs (Local Area Network) is analysed
- A traffic monitor is placed in each PoP
  - All incoming and outgoing traffic is monitored
  - Standard PCs used



Measured Properties

1. Network performance
2. User behaviour
3. Traffic self-similarity

1. Network Performance

- One way packet delay
  - Increased delay is a sign of higher traffic load
  - Packet timestamped on entering and exiting the network
  - Clocks synchronized with NTP
- Packet loss ratio
  - Increased PLR is sign of congestion
  - Sent packets which are not seen in expected destinations are considered lost
  - Possible traffic leak points are an issue!

2. User Behaviour

- Protocol composition
  - Transport and application layers (L4, L7)
- Traffic volume
  - Dependant of both user behaviour and network performance
- Flow analysis
  - Size distribution
  - Inter-arrival time distribution

Traffic model

3. Self-similarity

- If traffic volume as a function of time looks the same in all time-scales it is said to be self-similar, i.e. fractal-like
- Self-similarity is informative about users processes
  - if discovered, Poisson-model is not valid
  - → user think time or transfer size distribution is heavy-tailed, e.g. Pareto
- Hurst parameter
  - single scalar value
- Comparison between the real and simulated network traffics

Traffic Model (1)

- Flow based model
  - 5-tuple of source and destination IPs and ports and the protocol
  - Measured in the network
  - Modelled with TCP- and UDP-connections
- Traffic of the whole network is included in the model

Traffic Model (2)

- Traffic model consists of traffic generators
  - A traffic generator defines the rate at which flows are generated and the size of the flows
  - Each traffic generator has a unique triplet of
    - source and destination PoPs
      - Full mesh point-to-point connectivity
    - protocol
      - The number of protocols (P) can be reduced with aggregation
  - Network with N PoPs and P protocols
    - → PN(N-1) traffic generators

Initial Findings

- Traffic volume
  - Highly dependent on location
  - Highly dependent on time
    - Typical busy and slow hours during the day
    - Periodic traffic bursts
- Varying levels of self-similarity
  - Highly dependant on location
  - Heavy-tailed arrival processes also found
  - Time dependency not yet measured

Future Work

- Making the simulations
  - Comparison between the real and simulated networks
- Improving the traffic models
  - Increasing accuracy by studying packet level phenomena
  - Reducing the number of traffic generators with better clustering methods
  - Automating the creation
- More accurate timestamps
  - More work on NTP and Linux kernel tuning

Thank you!

Questions, comments?

timo.viipuri@hut.fi