Dynamic generation of filters for quality differentiation

IRoNet-seminar 8.1.2004

Mika Ilvesmäki

Contents

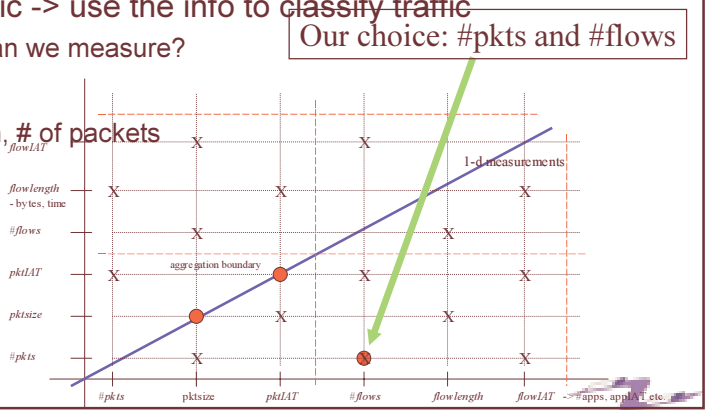
- Measurement based policy creation
  - Filters – definitions and generation
  - Application behaviour revealed with packet and flow measurements
- Measurement analysis
  - different choices for measurements and our choice (pkt/flow – space)
  - analysis choices of measurements
- Evaluating the classification results

General Introduction

- In a QoS environment we need to distinguish (at least) three traffic categories
  - Externally controlled traffic flow (video, voice)
  - TCP controlled traffic flows
  - Short duration traffic bursts
- Network, especially in DiffServ, may be allowed to perform the classification (DiffServ AF)
  - Can we do this based on measurements?

Measurement based policy creation

- Policy creation supports a QoS capable network
- Measure the network -> characterize the network/traffic -> use the info to classify traffic
  - But what can we measure?
- Packet
  - IAT, Length, # of packets
- Flow
  - IAT, length, # of flows



Design guidelines #1, #2 and #3

1. Do not associate port numbers to QoS classes (-> potentially 65535 classes)
   - Analyze traffic, get port number lists and bind the contents of the list to DiffServ Codepoints (DSCP), for instance.
     - Port number have nothing to do with QoS identification whereas DSCP is designed just for that
2. Do not imply policy within design
   - Use as value-neutral design as possible and leave room for freedom of choice
3. Preserve end to end principle: ”If possible do everything at the edges.”
   - Profiling and marking should be done and used at the edges of the network
     - although measurements may, of course, be done anywhere in the network

HELSINKI UNIVERSITY OF TECHNOLOGY Mika Ilvesmäki, Lic.Sc. (Tech.)

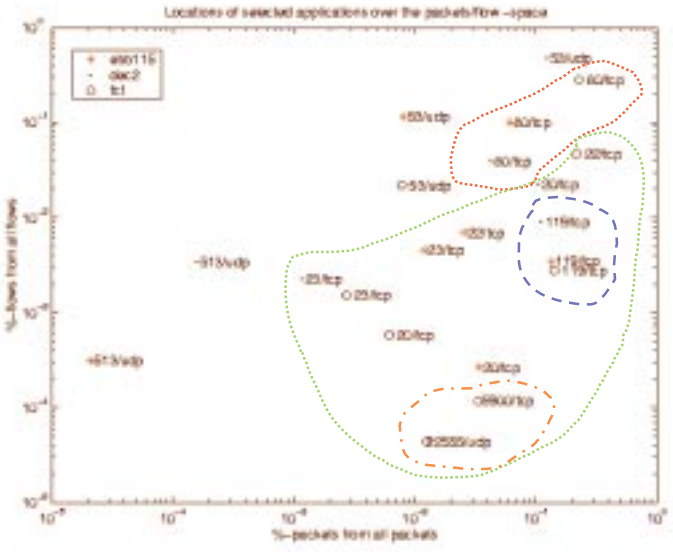
What is a filter?

- Our work uses TCP/UDP port numbers as filters
  - No packet content analysis
    - Too complicated to be used with every application (past and present)
  - Packet header information
    - layers 1 and 2 do not contain any information on packet content
    - layer 3 (IP) identifies the sending source and receiving destination the upper layer 4 protocol (TCP/UDP)
      - oversimplification: who sends packets where
    - layer 4 (UDP/TCP) identifies the port numbers used at source and destination
      - oversimplification: what application is used
      - source identifies the application that originates the packet and the destination tells us where the packets are headed

How are the filters generated?

- Based on packet and flow measurements
  - Results show that very simple analysis with packet and flow (5-tuple,60 second timeout) counts per Sport is enough to divide the traffic into two classes
    - Using other timeout values we can extend the division of traffic into several classes.
- And the classes (and their contents) make sense!

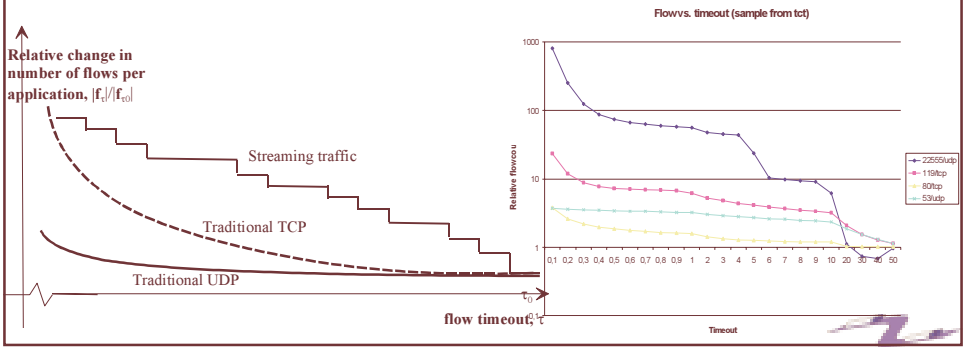
Packet/flow space



• Packet or flow data as standalone measurements do not reveal the application behavior
•Similar applications tend to position in the packet/flow –space the same way in different network environments
• Clustering should be (and partially has been) investigated more.

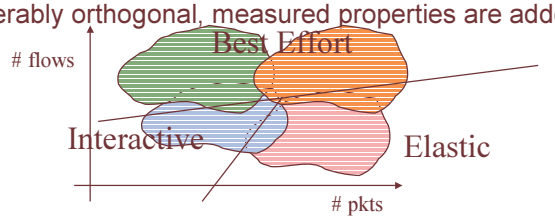
Changing the flow timeout

- Different types of traffic behave differently
  - Traffic behavior types may be found by observing the flow count with different flow timeout values



Summary on measurements

- Packet and flow counts indicate application characteristics when observed together
  - 2 or 3 classes may be identified
  - Varying the flow timeout and observing the flowcount gives a strong indication on the application behavior
- Our suggestion: Increase the dimensionality of the measurement analysis
  - Packet phenomena may be better described if new, preferably orthogonal, measured properties are added



Measurement analysis methods

- Measurements may be analyzed by simple statistics
  - averages, variances etc.
  - distribution modeling
- The measured/analyzed properties may also be sorted, or otherwise analyzed against
  - absolute boundaries (particular packet sizes, certain variance limits)
  - each other (all packets smaller/larger than the average packet size are classified/not classified)
- Multidimensional data may be clustered and classified
  - SOM, LVQ (if pre-classified samples are available) and other classification/cluster identification methods

Evaluation of the policy creation system

- Evaluate the network element (edge router)
  - Use of transmission capacity, architecture dependent router resources (connection setup / class, packet forwarding / class etc.)
  - RESULT: Edge routers will be able to handle the classification induced workload.
- Evaluate the effect on user
  - What applications are classified to priority
    - Relevance, application type, application count
    - Coherence of the application set

Application set analysis

- For each application group we have the packet and flow counts.
  - Observe the mean and the stdev and particularly the stdev/mean –ratio (less is better) for packet and flow data within class.
    - With different classification methods
      - No classification: $\approx 10^{1}$
      - Selected applications: $\approx 10^{-2}$
      - 2% of the apps with highest pkt/flw-ratio: $\approx 10^{-1}$
      - IRoNet-LVQ: $\approx 10^{-1} \ldots 10^{-2}$
  - Class content analysis shows that LVQ performance is comparable with selected apps.

HELSINKI UNIVERSITY OF TECHNOLOGY Mika Ilvesmäki, Lic.Sc. (Tech.)

Summary

- Measurements should be done on the packet level concentrating on the packet header infromation (and arrival information of the packet)
  - simple statistics are enough
- Analysis and further use of measurements is an upcoming field of research.
- Evaluation methods are beginning to emerge