

# Flow analysis of FUNET-data

IRoNet-seminar 17.1.2005

Mika Ilvesmäki

Networking laboratory



# Contents

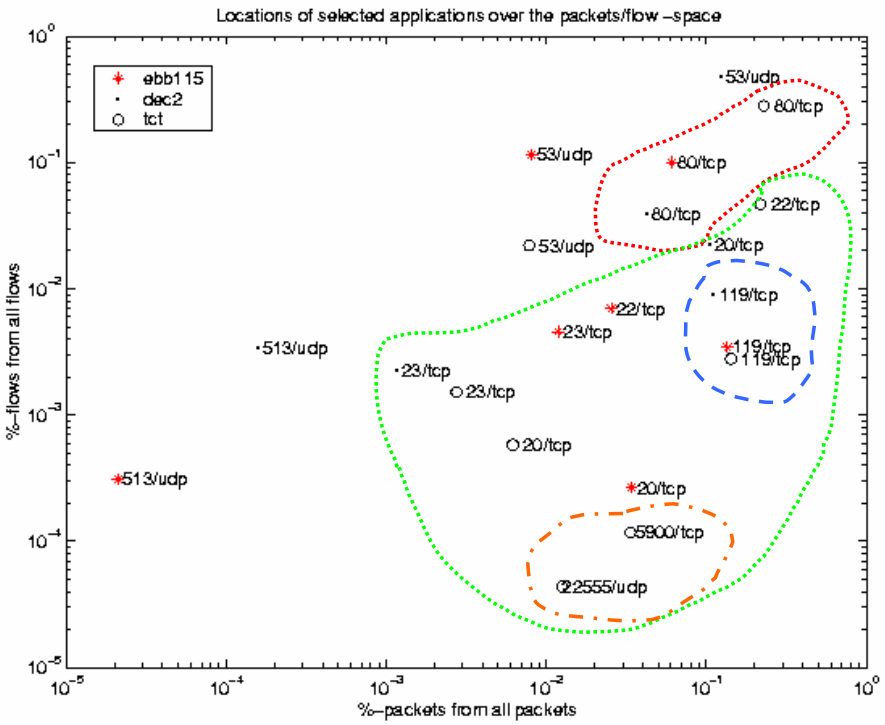
- Previous results
  - Confirmation of assumptions
- New preliminary results
  - Analysis of flow lifetime
  - Packet/byte/flowcount intensity of [Sport, Dport] -pairs
- Future research
  - Completing the analysis of “New preliminary results”
  - FlowIAT

# Introduction

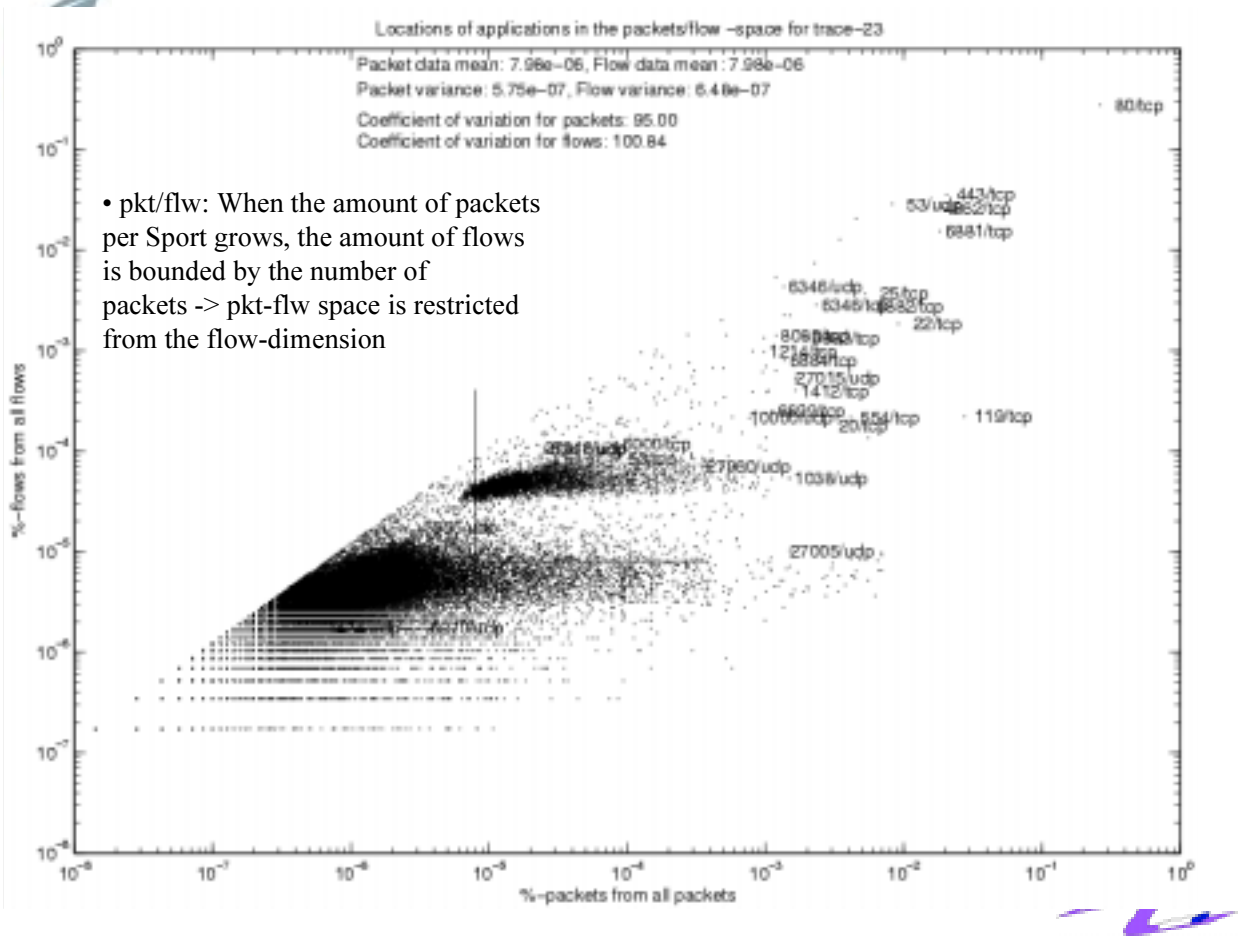
- Flow analysis has been used for traffic classification purposes
  - Results have shown that very simple analysis with packet and flow (5-tuple,60 second timeout) counts per Sport is enough to divide the traffic into at least two classes
- However, limited access to traffic measurement data has casted a shadow on our results



# Packet/flow space



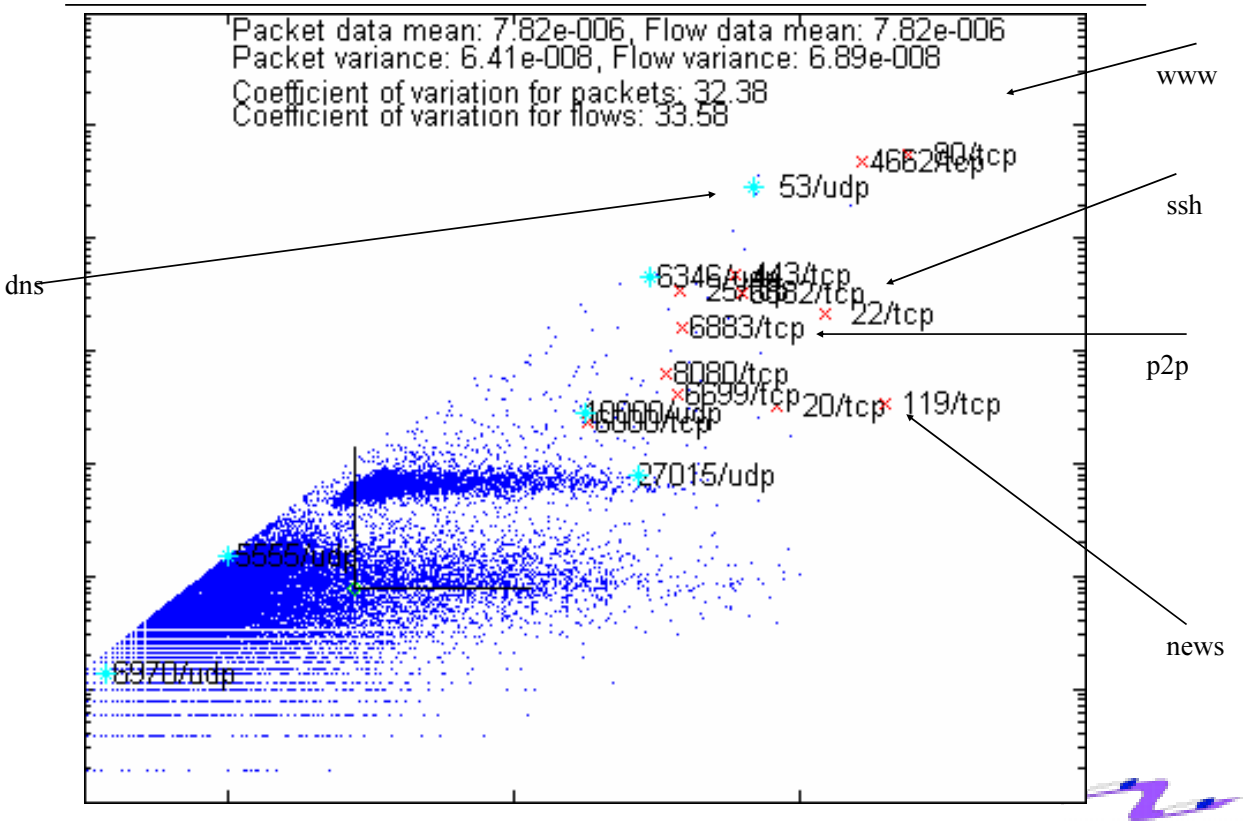

• Packet or flow data as standalone measurements do not reveal the application behavior
•Similar applications tend to position in the packet/flow –space the same way in different network environments
• Clustering should be (and partially has been) investigated more.




• pkt/flw: When the amount of packets per Sport grows, the amount of flows is bounded by the number of packets -> pkt-flw space is restricted from the flow-dimension

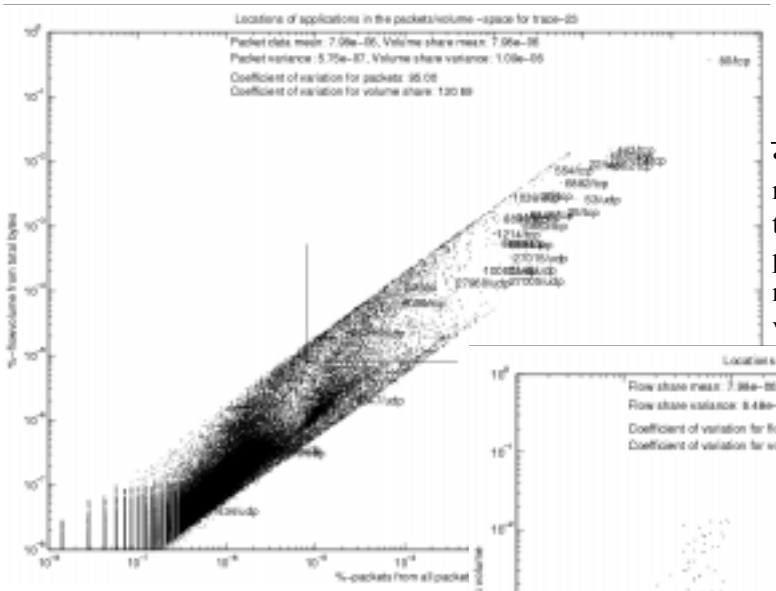
# App. movement in pkt/flw -space

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



# Previous assumption confirmed

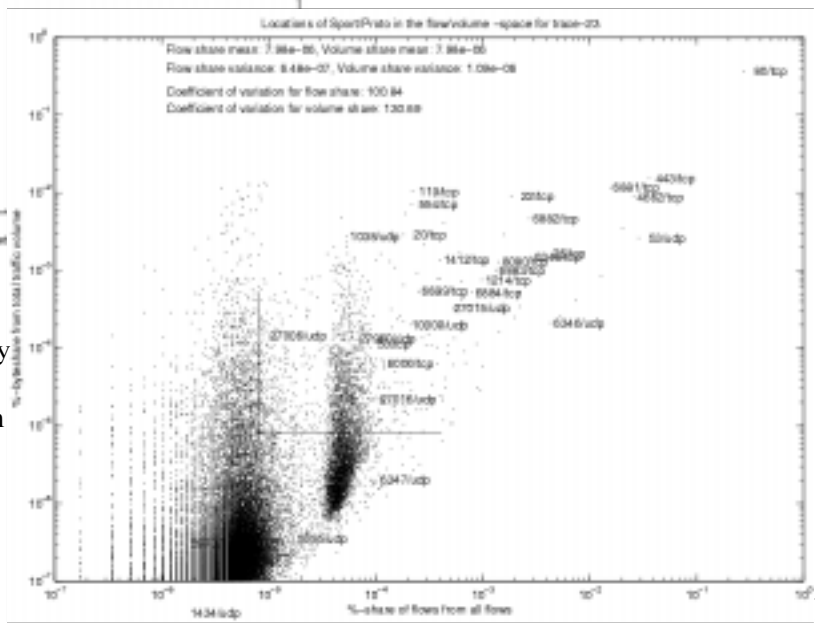
- Similar applications tend to position in the packet/flow – space the same way in different network environments
  - Changes occur according to application usage
  - Application movement analysis in the pipeline -> future research!!



*Pkt-bytevolume/Sport and flw/bytevolume/Sport*

Mika Ilvesmäki, Lic.Sc. (Tech.)

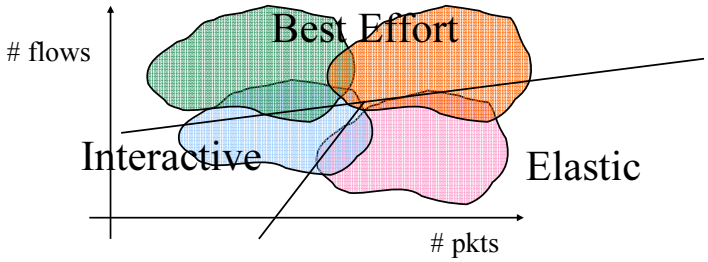
• pkt/bytevolume: Packets have a minimum and maximum size, therefore when the amount of packets per Sport grows there is always a minimum and a maximum for the byte volume



• flw/bytevolume: Flows always carry packets (with min size). Flows, however are not limited in how much packets they are allowed to carry. Therefore when the flowcount per Sport grows there is always a minimum for the byte volume

# Summary on measurements

- Packet and flow counts indicate application characteristics when observed together
  - 2 or 3 classes may be identified
- Future research: Introduce other packet and flow properties for analysis!!




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

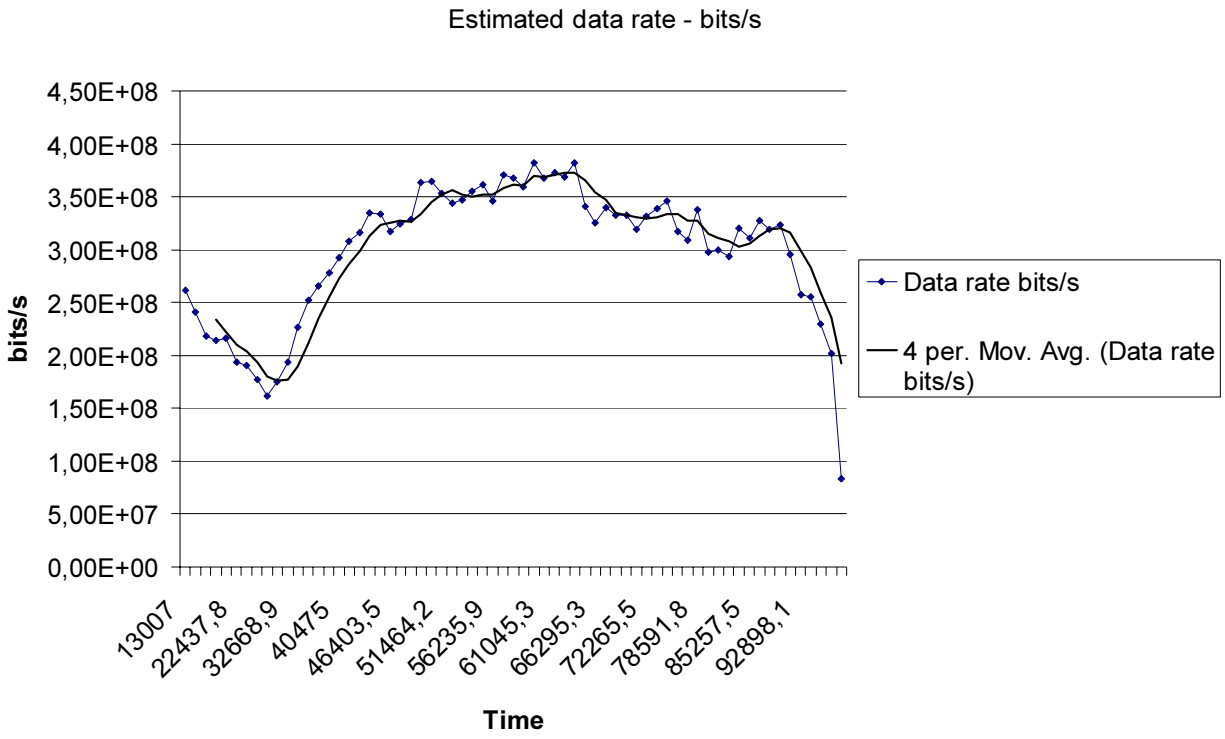
# New measurements

- New data from FUNET
  - Trace was captured on November 30th, 2004 starting at midnight (UTC) and ending 26 hours later. Flows were formed based on 64 second timeout and 5-tuple flow granularity.
    - Due to file size limitations the trace was divided into 65 parts. The whole trace contains 3499 Gbytes of data in 4,7Gpackets in 350 million flows.
  - The trace parts are on the average 1486 seconds (almost 25 minutes) in length. One trace part contains 53Gbytes of data, transmitted 298 Mbit/s on average. One trace contains, on average, 72 million packets and 5 million flows.

# Basic flow data

- TCP: 60% of flows contain just 1 packet
  - 96% of packets in flows with more than 1 pkt
    - 35 pkts/flow
- UDP: 88% of flows contain just 1 packet
  - 85% of packets in flows with more than 1 pkt
    - 50 pkts/flow
- Other proto: 70% of flows contain just 1 pkt
  - 90% of packets in flows with more than 1 pkt
    - 50 pkts/flow

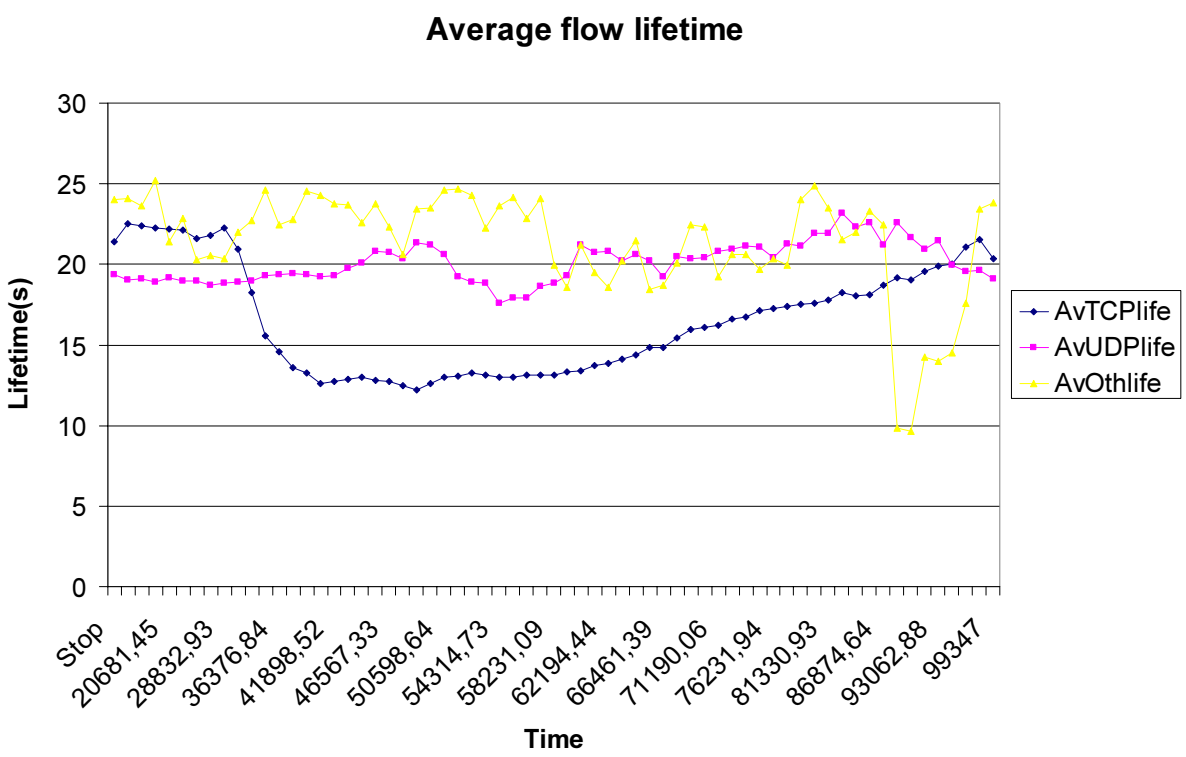
HELSINKI UNIVERSITY OF TECHNOLOGY
Mika Ilvesmäki, Lic.Sc. (Tech.)

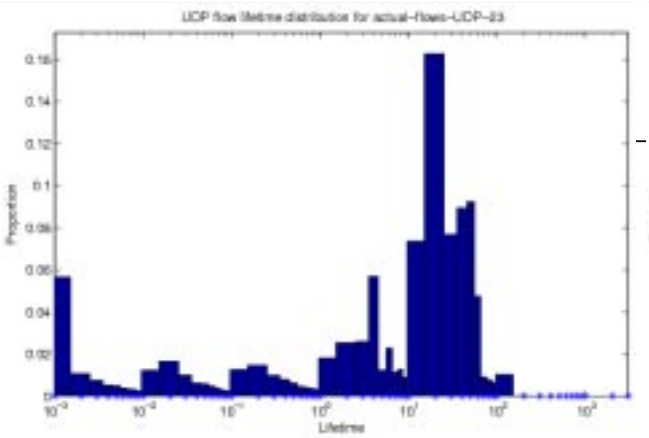


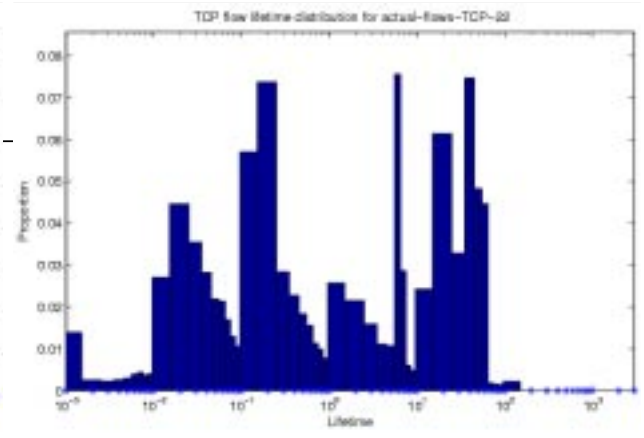
# Flow lifetime distributions for TCP and UDP

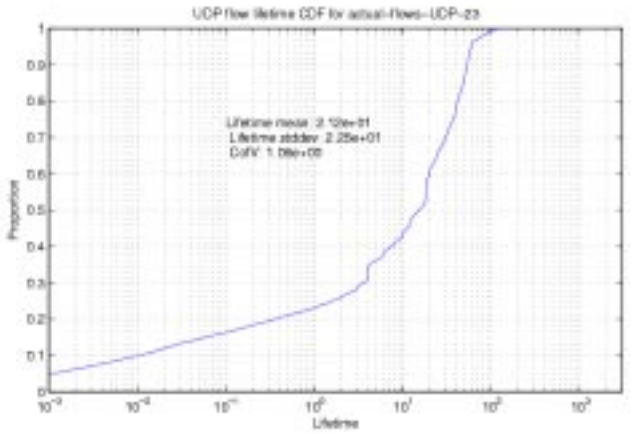
- TCP flow lifetime depends up on time of day (more activity -> shorter flows)
- Absence of long flows!
  - TCP: 15 s/flow (max 237s)
  - UDP: 20 s/flow (max 230s)
    - Only prelim Flow IAT analysis done
  - Suspected reason for flow shortness: TCP timeouts (due to packet loss)
- Interesting stairlike behavior in lifetime decades
  - No explanation yet.

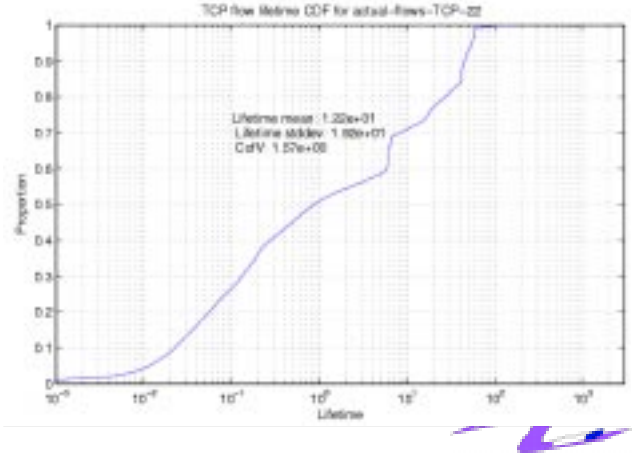
HELSINKI UNIVERSITY OF TECHNOLOGY
Mika Ilvesmäki, Lic.Sc. (Tech.)





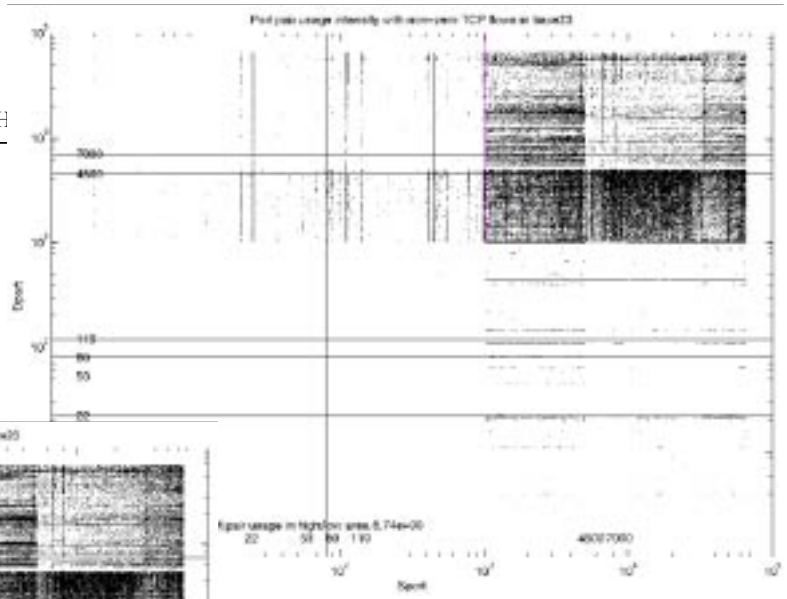



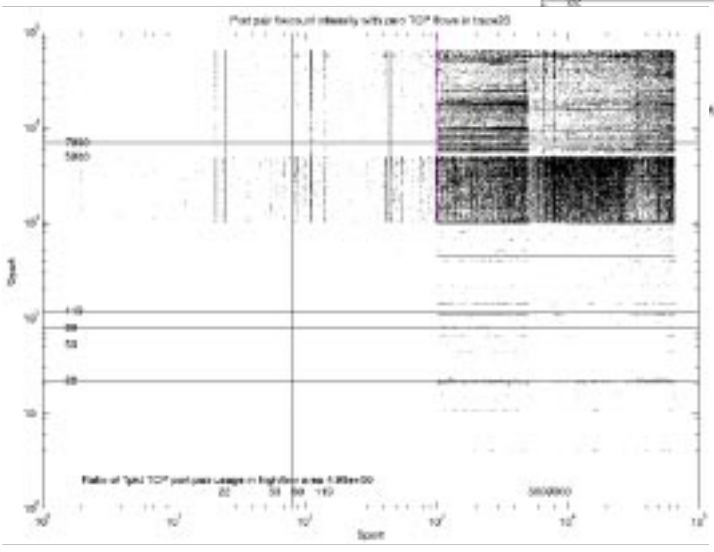




## Sport/Dport –use/TCP

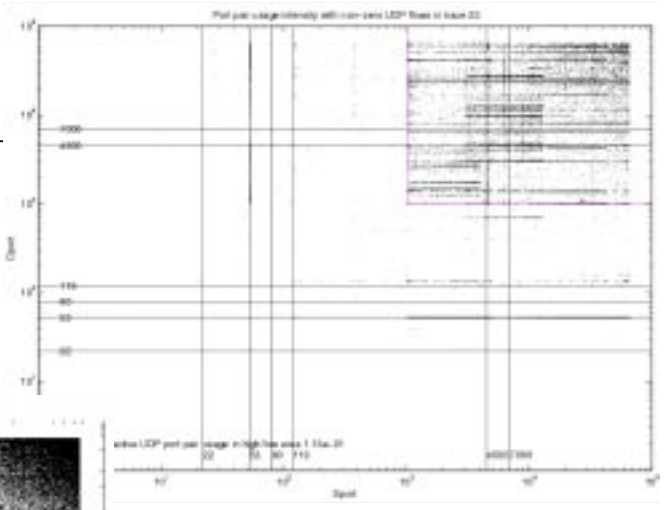
HELSINKI UNIVERSITY OF TECH

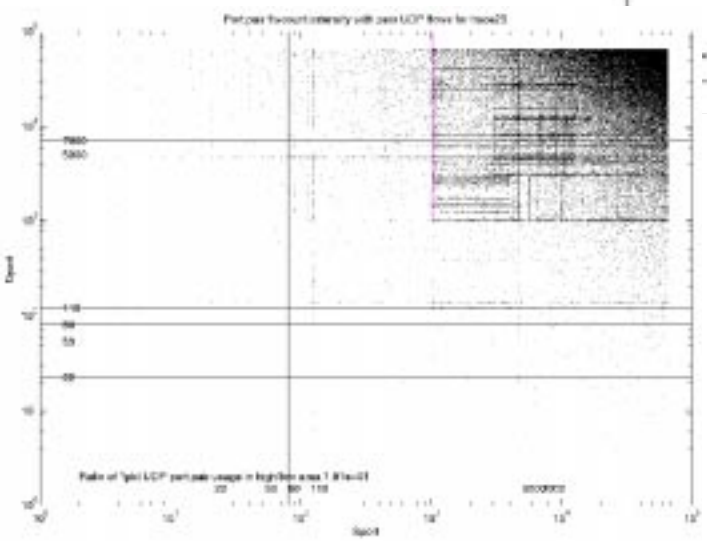


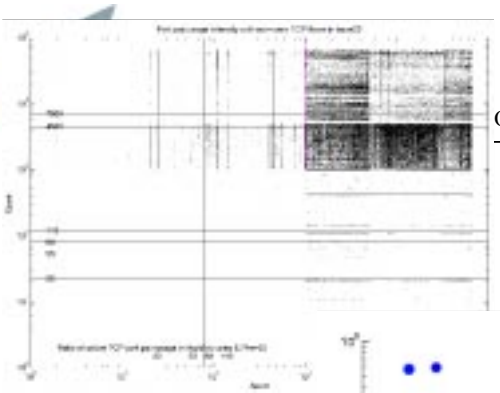




# Sport/Dport –use/UDP

HELSINKI UNIVERSITY OF TECHNOLOGY





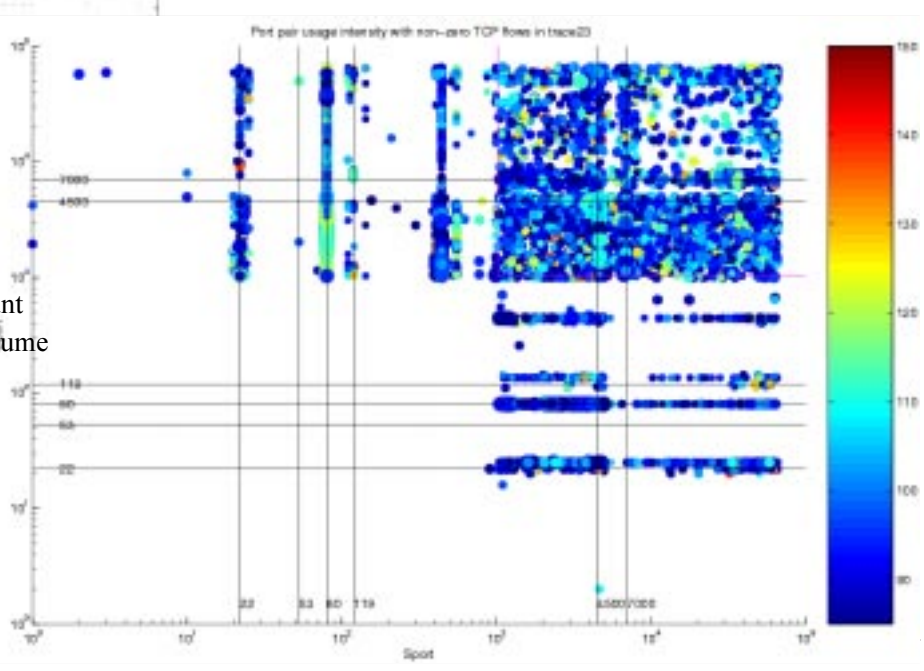


OLOGY

# Intensity (flowcount and bytevolume / pair)

Mika Ilvesmäki, Lic.Sc. (Tech.)

The pairs in the top 5% of the bytevolume shown:

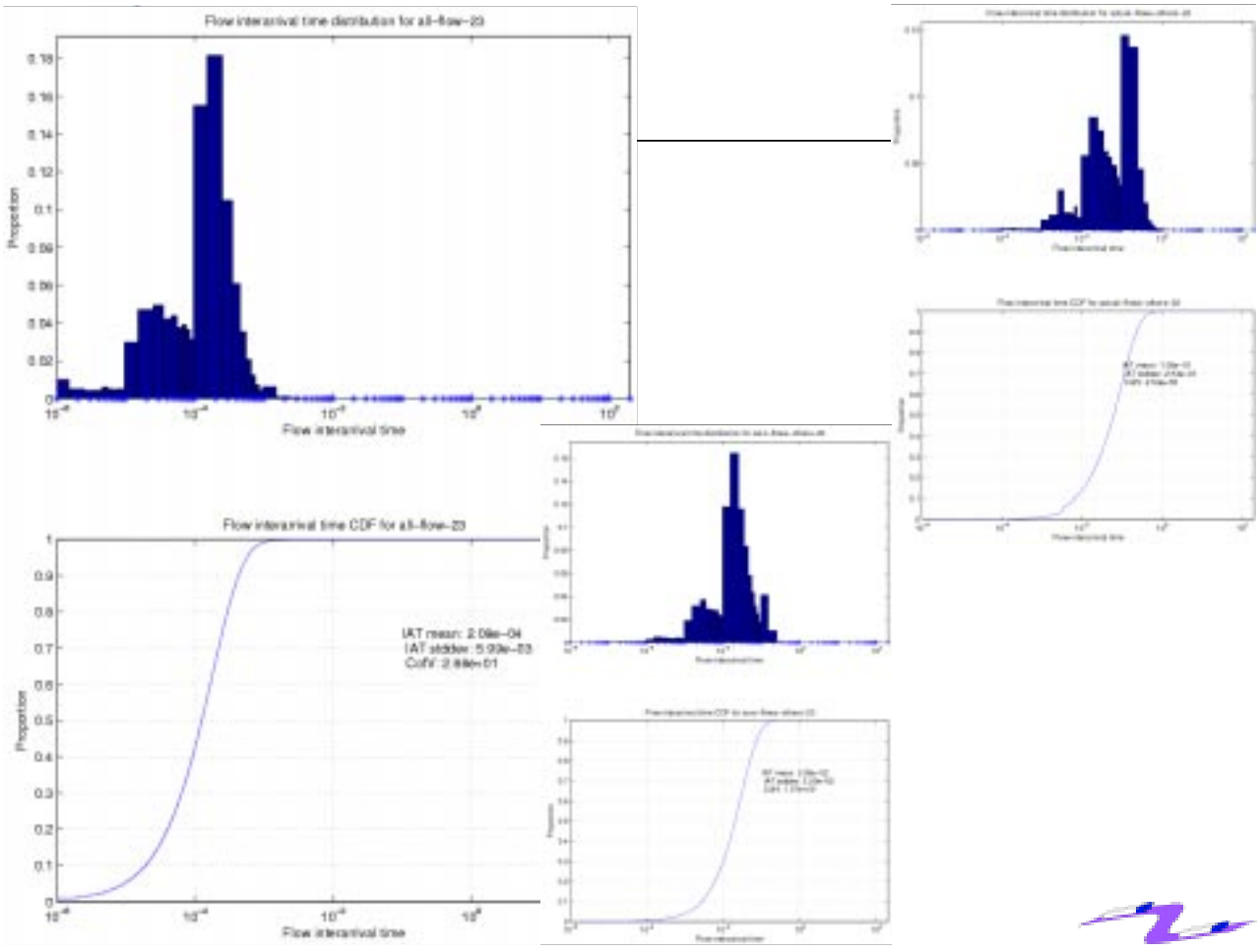


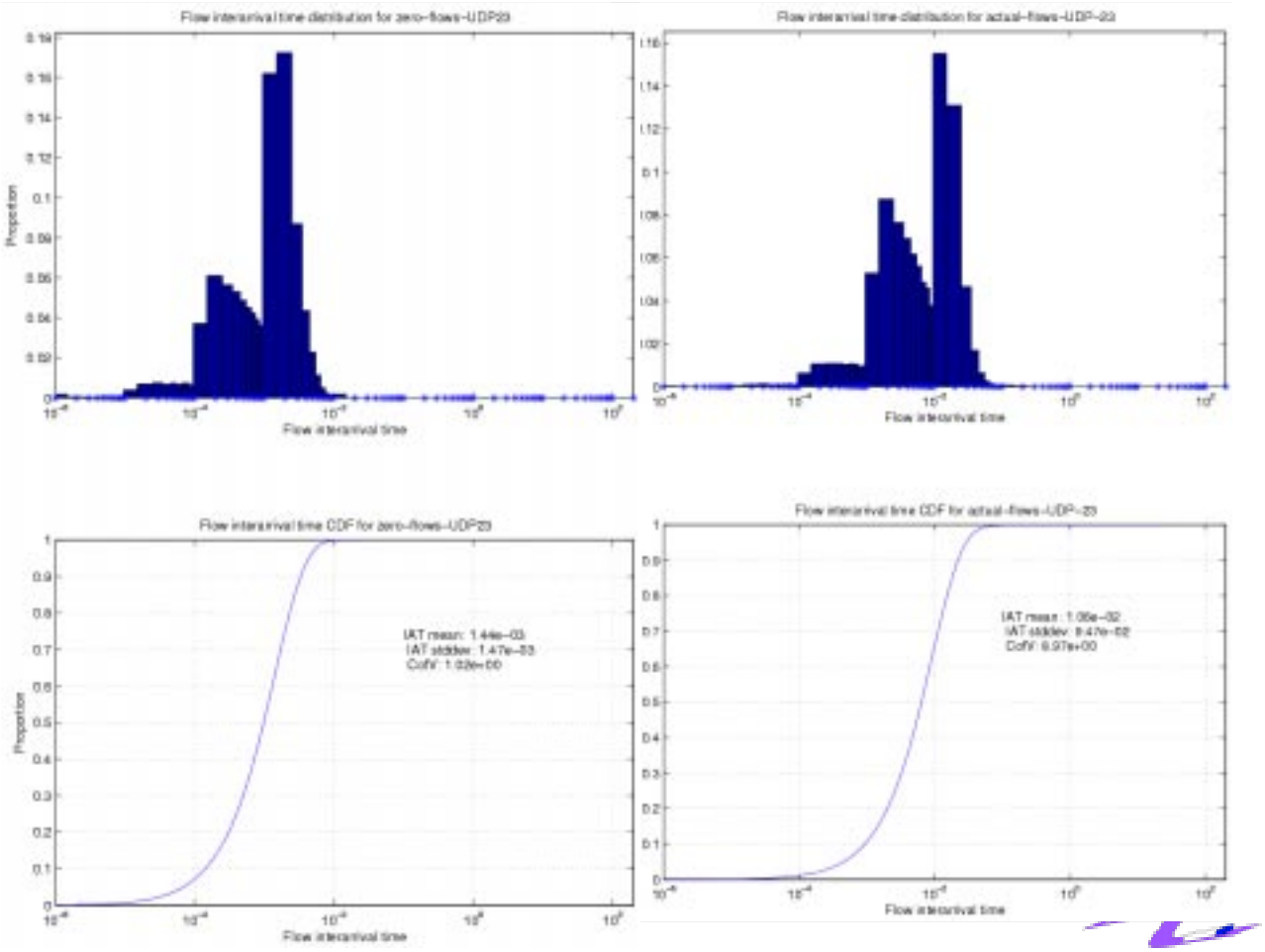

- dotsize represents flowcount
- dotcolour represents bytevolume

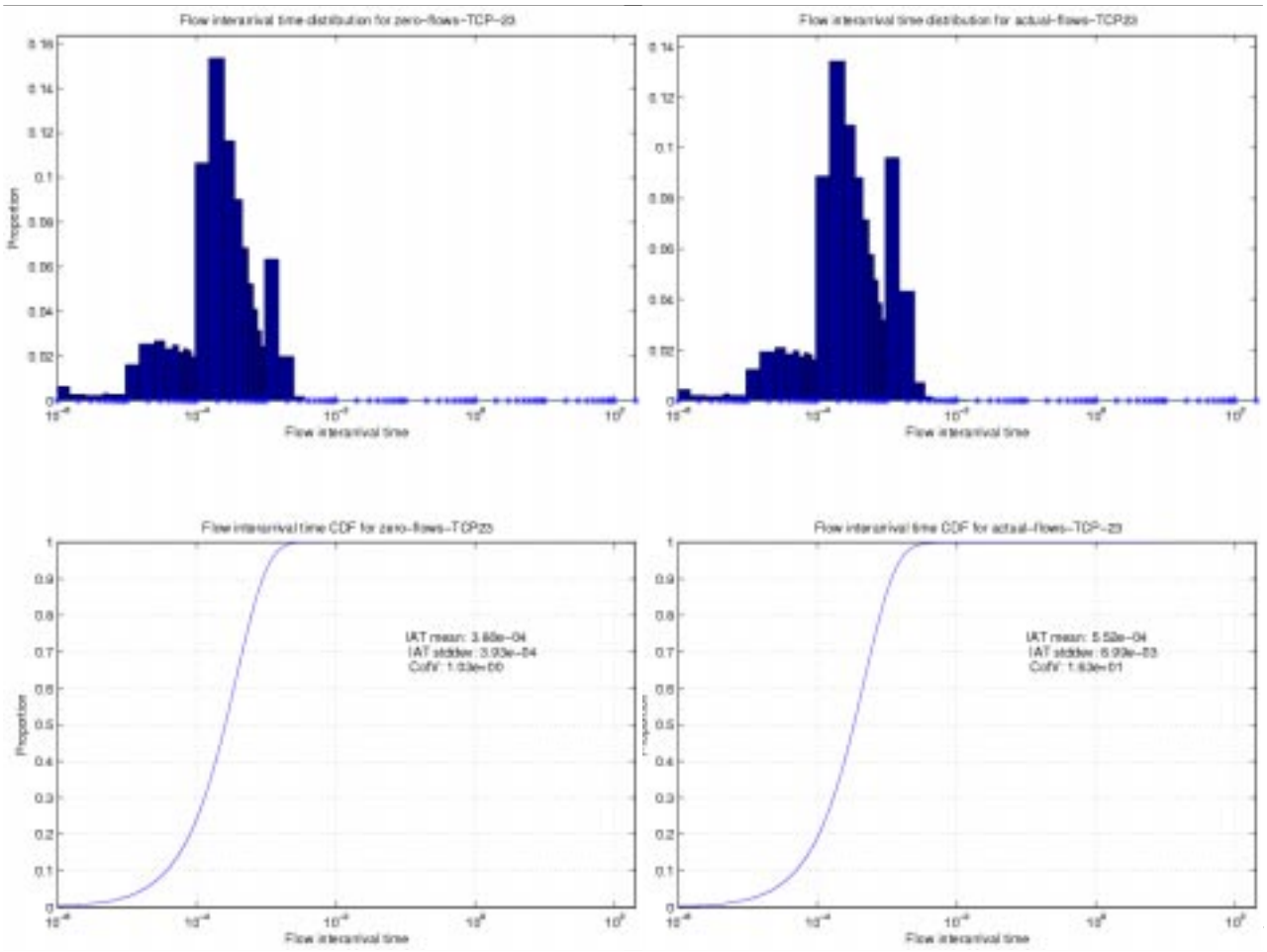
# More results to be published

- Technical report with more and detailed packet and flow data to be released in February/March-2005.

- Stay tuned!




Flow interarrival time distribution for zero-flows-UDP23
Flow interarrival time distribution for actual-flows-UDP-23
Proportion
Flow interarrival time
Flow interarrival time CDF for zero-flows-UDP23
IAT mean: 1.44e-03
IAT stddev: 1.47e-03
CofV: 1.02e+00
Flow interarrival time CDF for actual-flows-UDP-23
IAT mean: 1.06e-02
IAT stddev: 9.47e-02
CofV: 8.97e+00


Flow interarrival time distribution for zero-flows-TCP-23
Flow interarrival time distribution for actual-flows-TCP23
Proportion
Flow interarrival time
Flow interarrival time CDF for zero-flows-TCP23
IAT mean: 3.60e-04
IAT stddev: 3.93e-04
CofV: 1.03e+00
Flow interarrival time CDF for actual-flows-TCP-23
IAT mean: 5.52e-04
IAT stddev: 8.93e-03
CofV: 1.63e+01

# Future research / post-IRoNet

- Flow (and packet) inter-arrival-time study
  - One paper published in IRoNet, results need to be confirmed with new data
  - Incorporating results to traffic classification
- more [Sport,Dport]- usage studies
  - Determine usefulness for differentiating/detecting traffic
- more Sport / Dport activity studies
  - Application arrivals (and departures)