Research Academic Computer Technology Institute
University of Patras

Analyzing Traffic across the Greek School Network

Costas Kattirtzis, Emmanuel Varvarigos, Kyriakos Vlachos,
University of Patras & Research Academic Computer Technology Institute

George Stathakopoulos and Michael Paraskevas
Research Academic Computer Technology Institute

LANMAN 2005, 14th IEEE Workshop on Local and Metropolitan Area Networks,
18-21 September 2005, Chania, Crete, Greece

Introduction

- Internet is growing dramatically.
- Very complex patterns to model the Network Traffic.
- Studies in LAN and WAN have been made since the early 80s.
- Today's findings lead us to the conclusion that
  - Ethernet traffic is statistically self-similar
  - Poisson assumption is valid in special cases
- Recent studies on Peer-to-Peer traffic mainly by Karagiannis et. al have been made.

Introduction

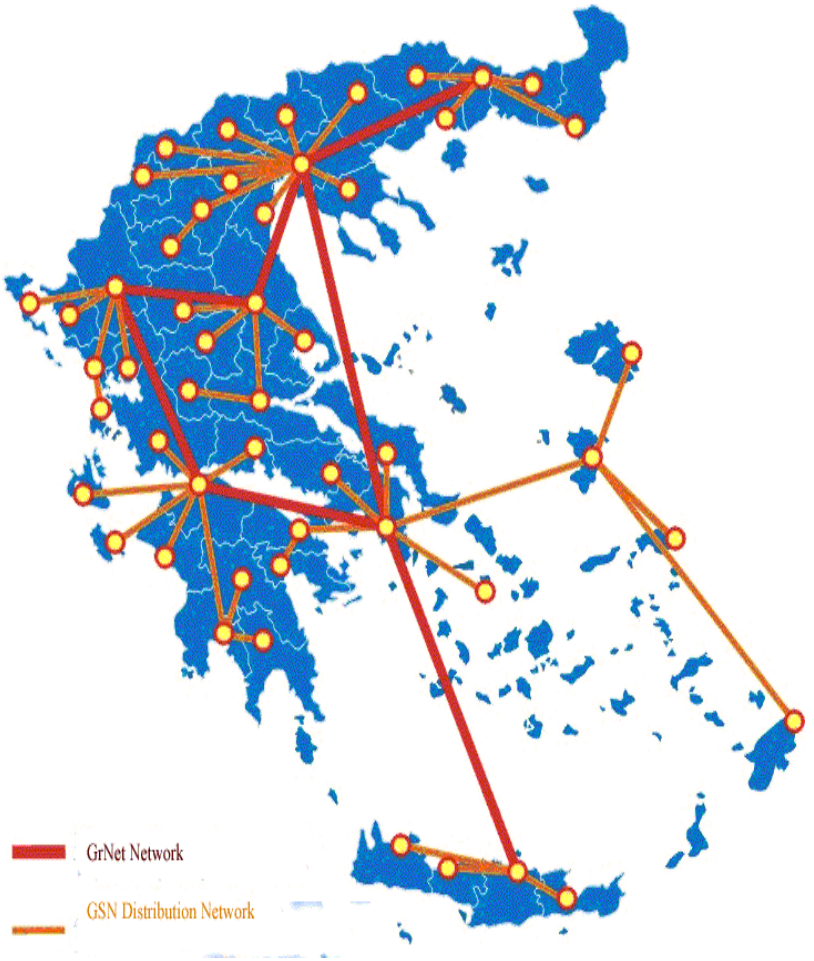
- In this paper we present a study of traffic patterns on the Greek School Network
- We studied in the monitored network
  - the behavior of flows
  - the behavior of the packets
  - the use of each protocol
  - the use of each well known application
  - The use of Peer-to-Peer services
  - The traffic locality phenomenon
- Benefits
  - Understand the impact of network changes and services
  - Improve network usage and application performance
  - Reduce IP service and application costs
  - Optimize network costs
  - Understand the Impact of P2P applications
  - Background to the administrators for
    - dimensioning the network
    - congestion control
    - network management

Overview

- Network Architecture
- Measurement Methodology
- Traffic Statistics
  - Service Analysis
  - Protocol Analysis
  - Flow Analysis
  - Packet Size Analysis
- Traffic locality
- Peer-to-Peer Services
- Conclusions

Greek School Network Architecture

- Nationwide network that spans across Greece. Connects all schools of primary and secondary education including administrator offices.
- Hierarchically structured into three layers.
  - The Backbone network
  - The Distribution Network
  - The Access Network



Overview

- Network Architecture
- Measurement Methodology
- Traffic Statistics
  - Service Analysis
  - Protocol Analysis
  - Flow Analysis
  - Packet Size Analysis
- Traffic locality
- Peer-to-Peer Services
- Conclusions

Measurement Methodology

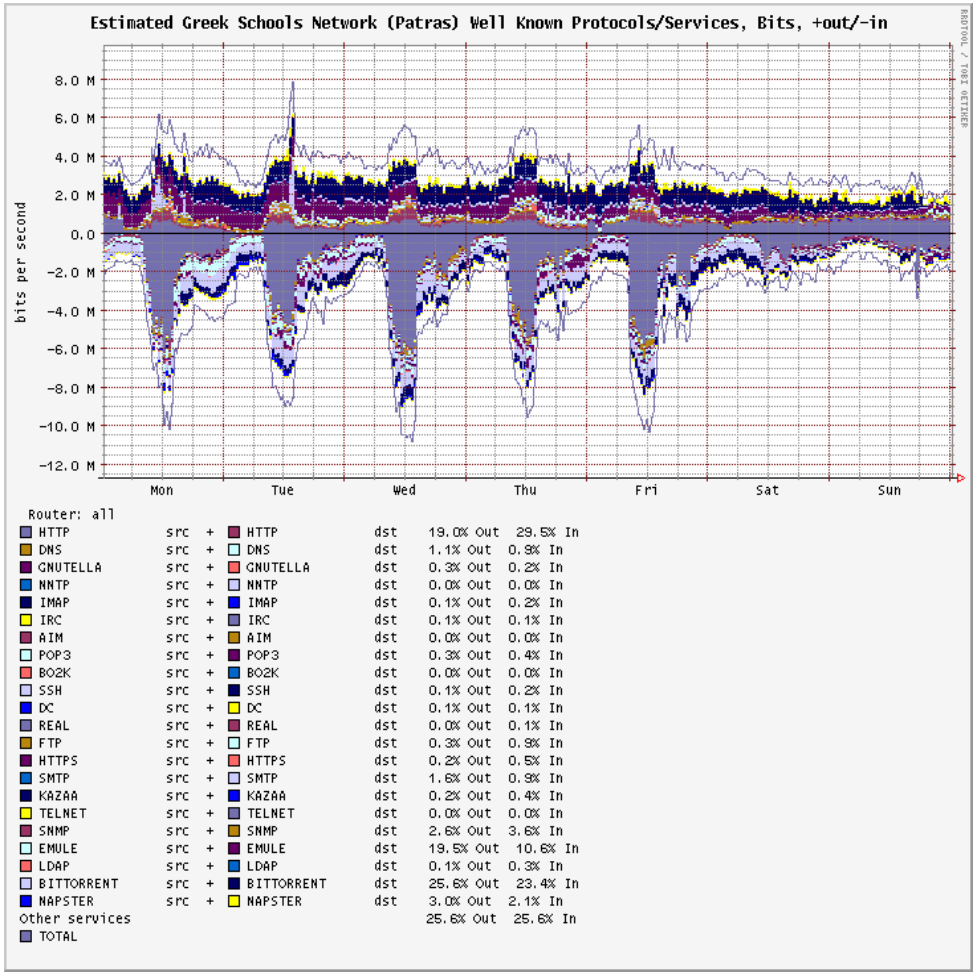
- All the measurements took place in the PATRAS prefecture from October 24 00:00:00 GMT+02:00 2004 to March 18 23:30:00 GMT+02:00 2005.

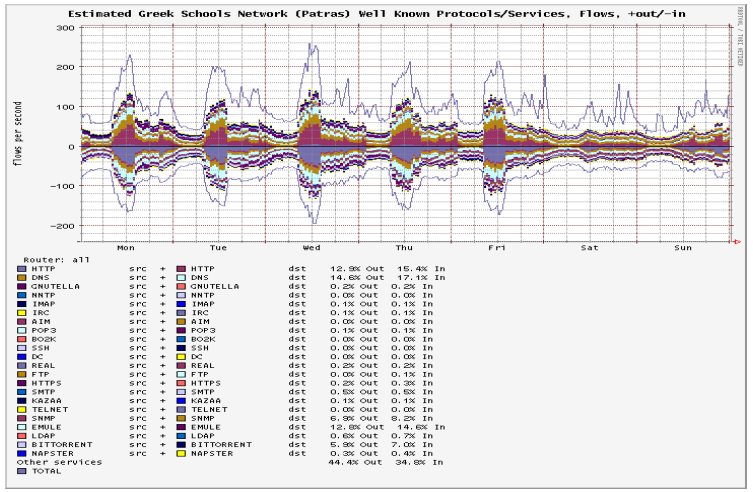
- Monitoring System
  - Cisco NetFlow
    - In terms of NetFlow, flow is defined by Seven Unique Keys:
      - source IP address
      - destination IP address
      - source port number
      - destination port number
      - layer 3 protocol type
      - TOS (Type Of Service) byte and
      - Input logical interface
  - FlowScan
  - cflowd
  - RRDtool

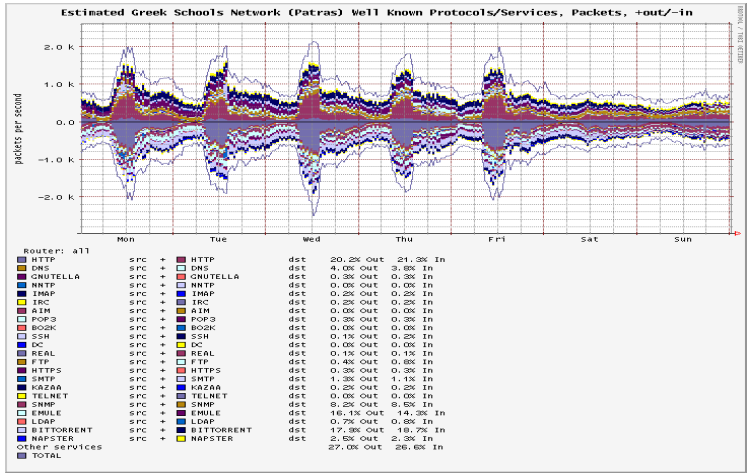
Overview

- Network Architecture
- Measurement Methodology
- Traffic Statistics
  - Service Analysis
  - Protocol Analysis
  - Flow Analysis
  - Packet Size Analysis
- Traffic locality
- Peer-to-Peer Services
- Conclusions

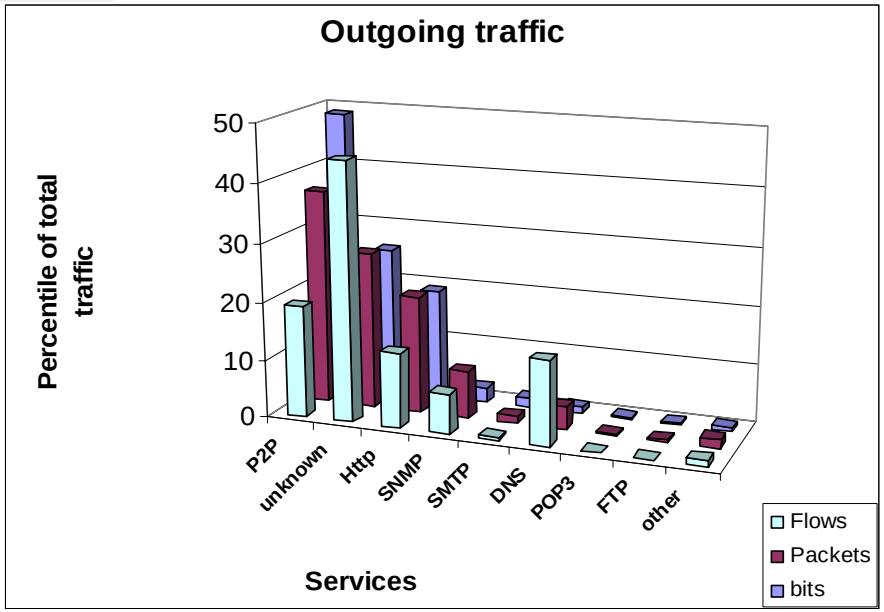
Traffic Statistics - Services

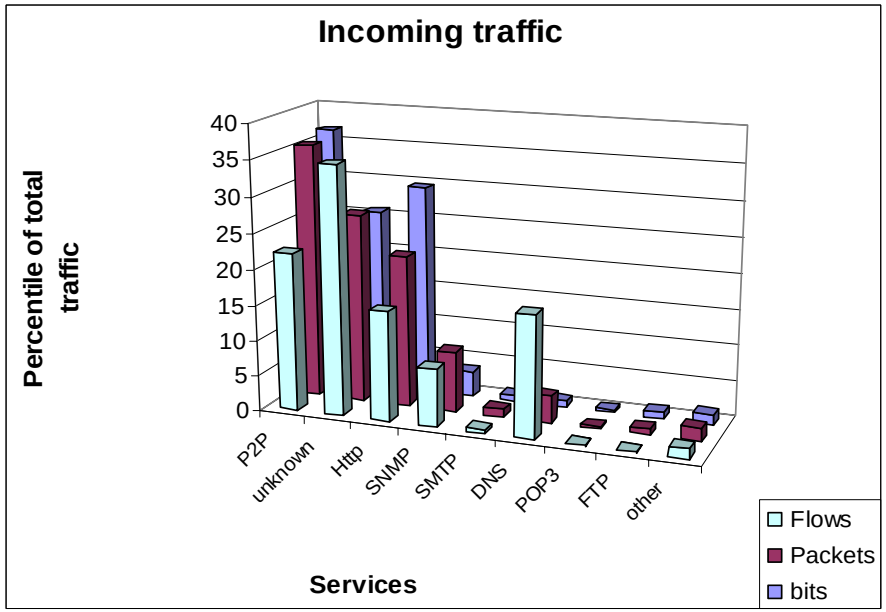






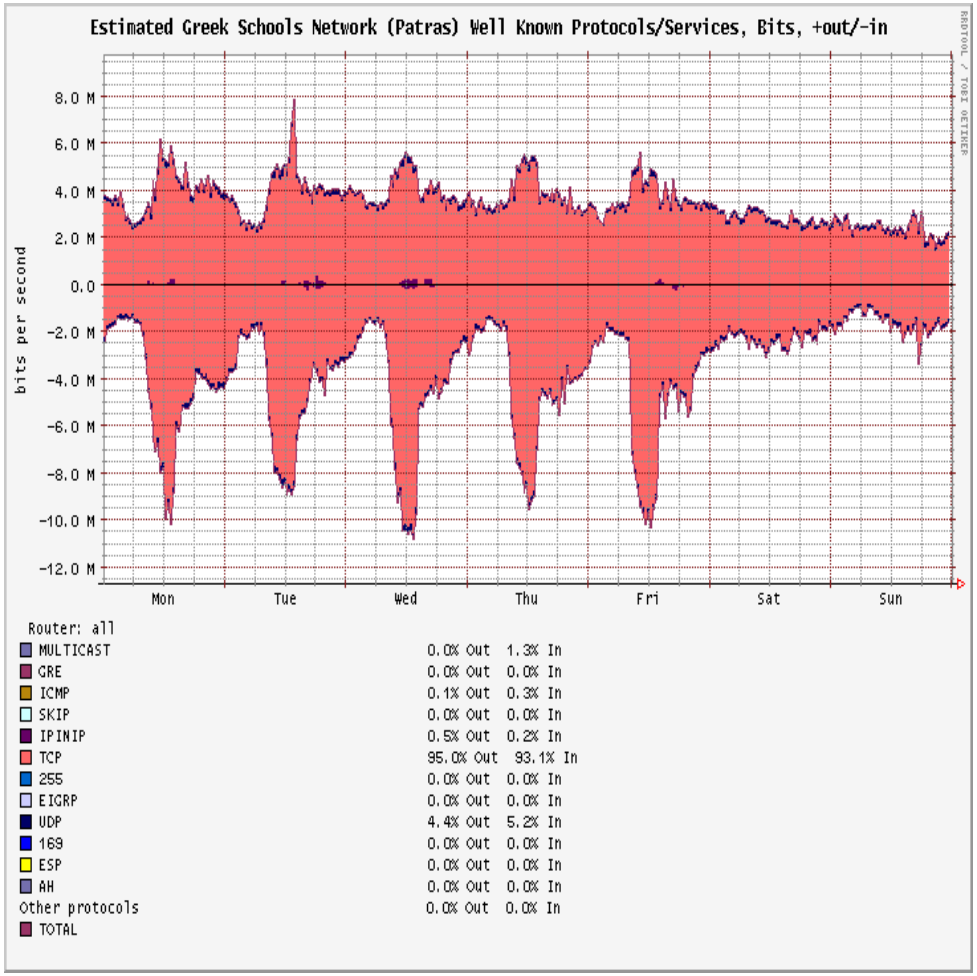
Traffic Statistics - Services

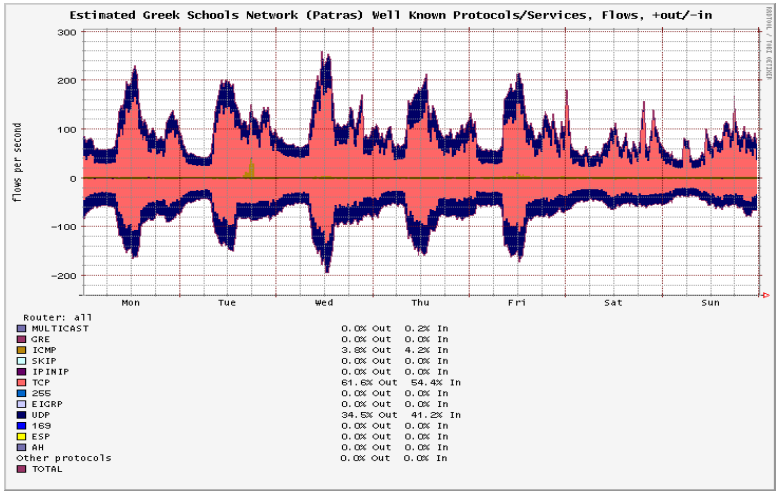


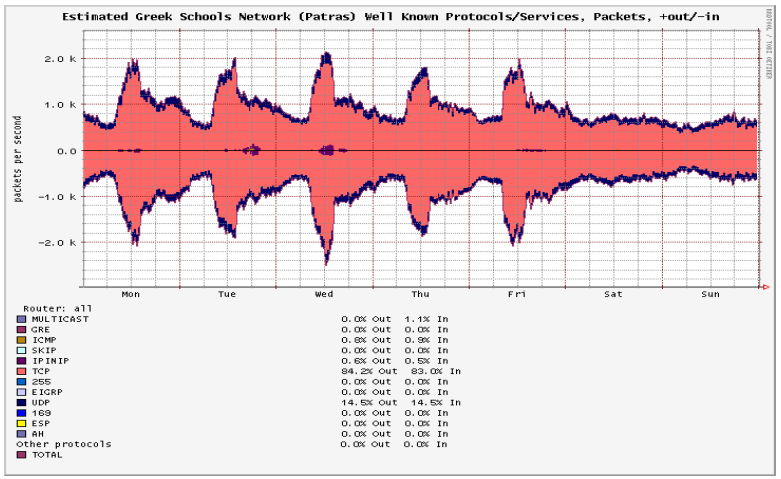


- Outgoing traffic in term of bytes
  - 50% is P2P
  - 19% is HTTP
  - 25.6% is unknown
- Incoming traffic in term of bytes
  - 37% is P2P
  - 30% is HTTP
  - 25.6% is unknown
- DNS and SNMP use UDP
  - Large fraction of the flows, small fraction of the packets and an even smaller fraction of the bytes transferred
- HTTP (web) application
  - The profile of its daily load distribution fits closely the corresponding profile of the TCP protocol.

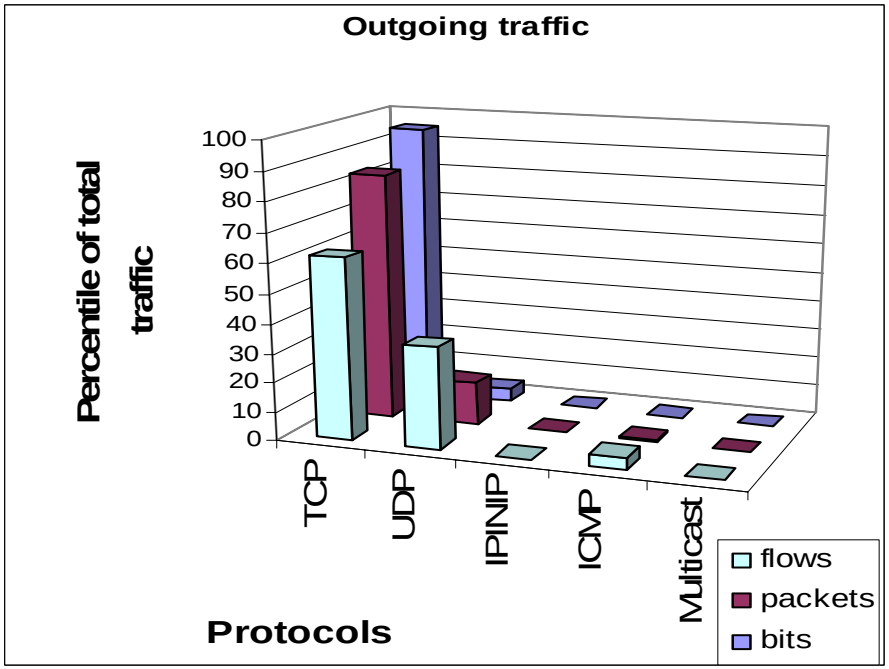
Traffic Statistics - Protocols

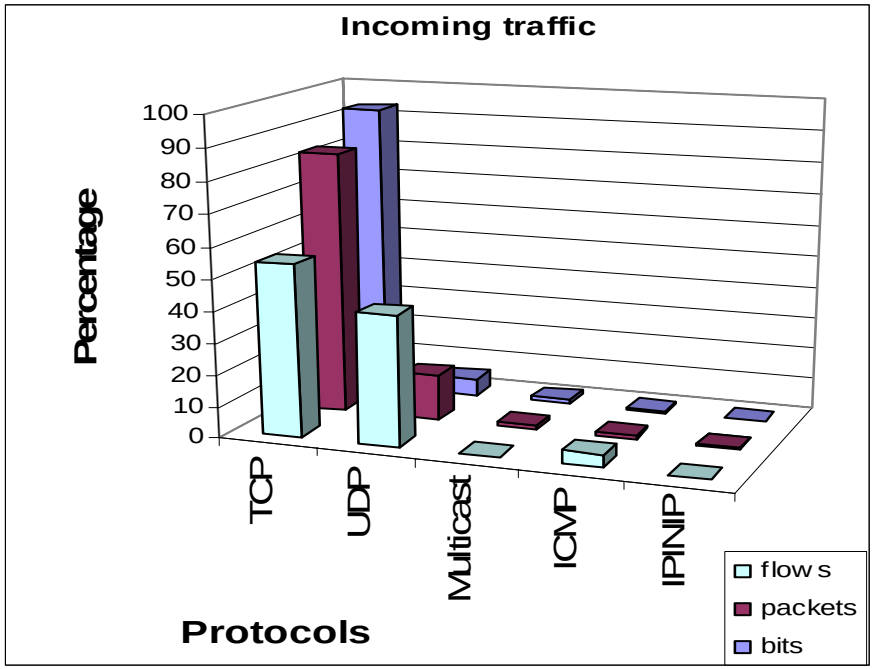






Traffic Statistics - Protocols



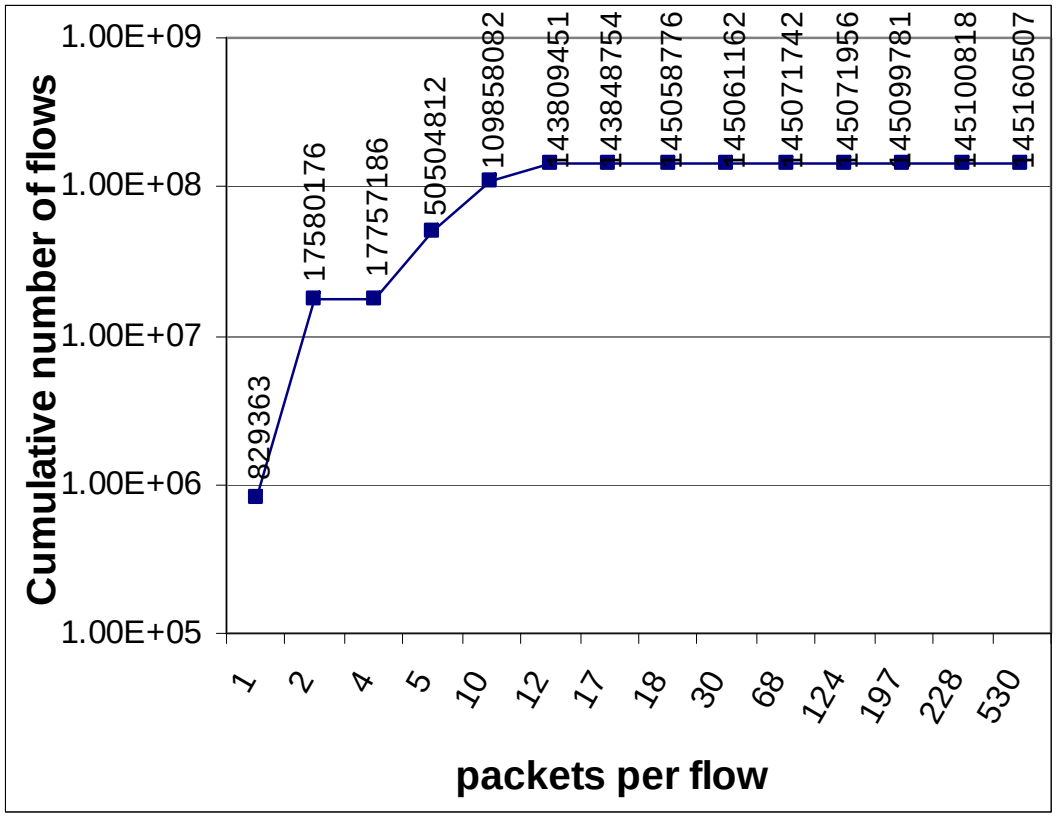


	Outgoing traffic			Incoming traffic		
Protocols	Bytes	Flows	Packets	Bytes	Flows	Packets
TCP	95%	61.6%	84,2%	93.1%	54.4%	83%
UDP	4,4	34,5	14,5	5,2	41,2	14,5

- The other IP protocols individually make up a negligible percentage of the overall traffic

- The size of the incoming packets is much larger than the size of the outgoing packets.
- TCP uses more and larger packets per flow than UDP

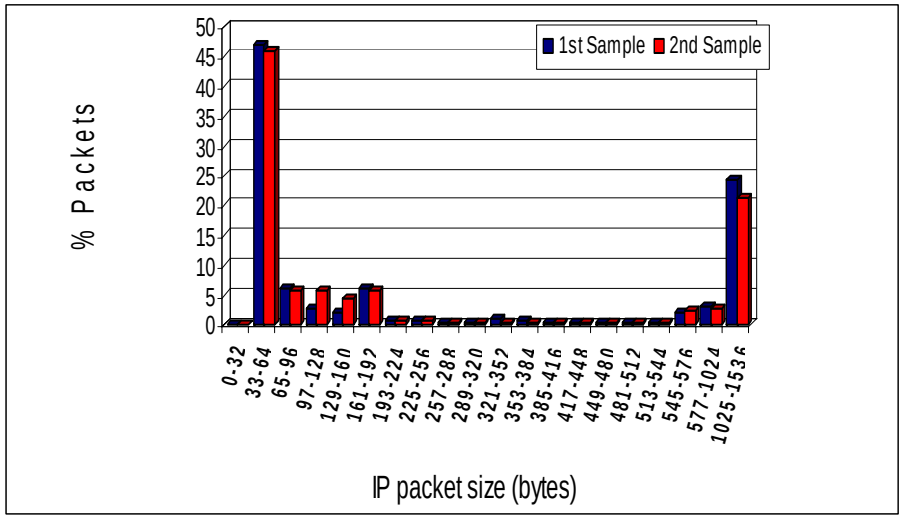
Traffic Statistics – Flow Analysis

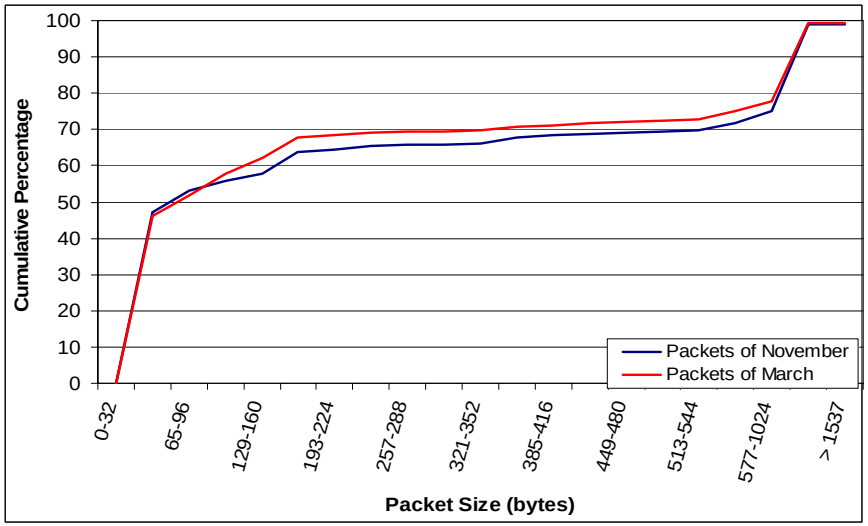


- 87% of the flows carry 5-12 packets
- The majority of the flows last 6 - 6.5 sec.
- Data transfers*
  - interactive: TCP-telnet, ICMP, UDP-NTP
  - transaction oriented: TCP-FTP, TCP-SMTP
  - bulk data transfer: TCP-FTPD, TCP-WWW

- A cross-check of the findings of k. Claffy et al. at “Traffic Characteristics of the T1 NSFNET Backbone”.

Traffic Statistics – Packet Size Analysis





- Dual-modal pattern
- Predominance of small-sized packets caused
  - by TCP control segments

  and
  - by HTTP application

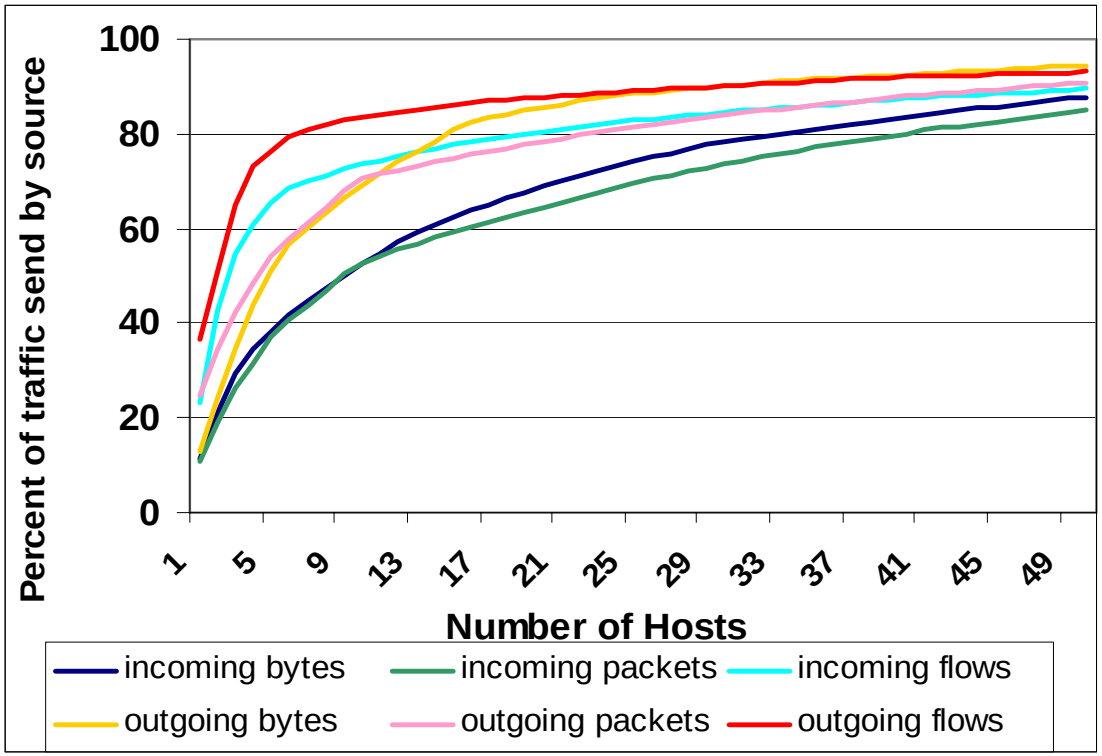
- Large size packets caused
  - By Ethernet full size packets

  and
  - By p2p applications

Overview

- Network Architecture
- Measurement Methodology
- Traffic Statistics
  - Service Analysis
  - Protocol Analysis
  - Flow Analysis
  - Packet Size Analysis
- Traffic locality
- Peer-to-Peer Services
- Conclusions

Traffic Statistics – Traffic Locality



- Outgoing traffic: The 50 most busy sources (of the 6188) in a 5-minute sample, are responsible for
  - 94.5% of the bytes
  - 93.1% of the flows
  - 90.9% of the packets.
- Incoming traffic: The same users:
  - 76.6% of the bytes
  - 77.5% of the flows
  - 52.5% of the packets.
- The same results were observed in the 250 minutes samples.

Overview

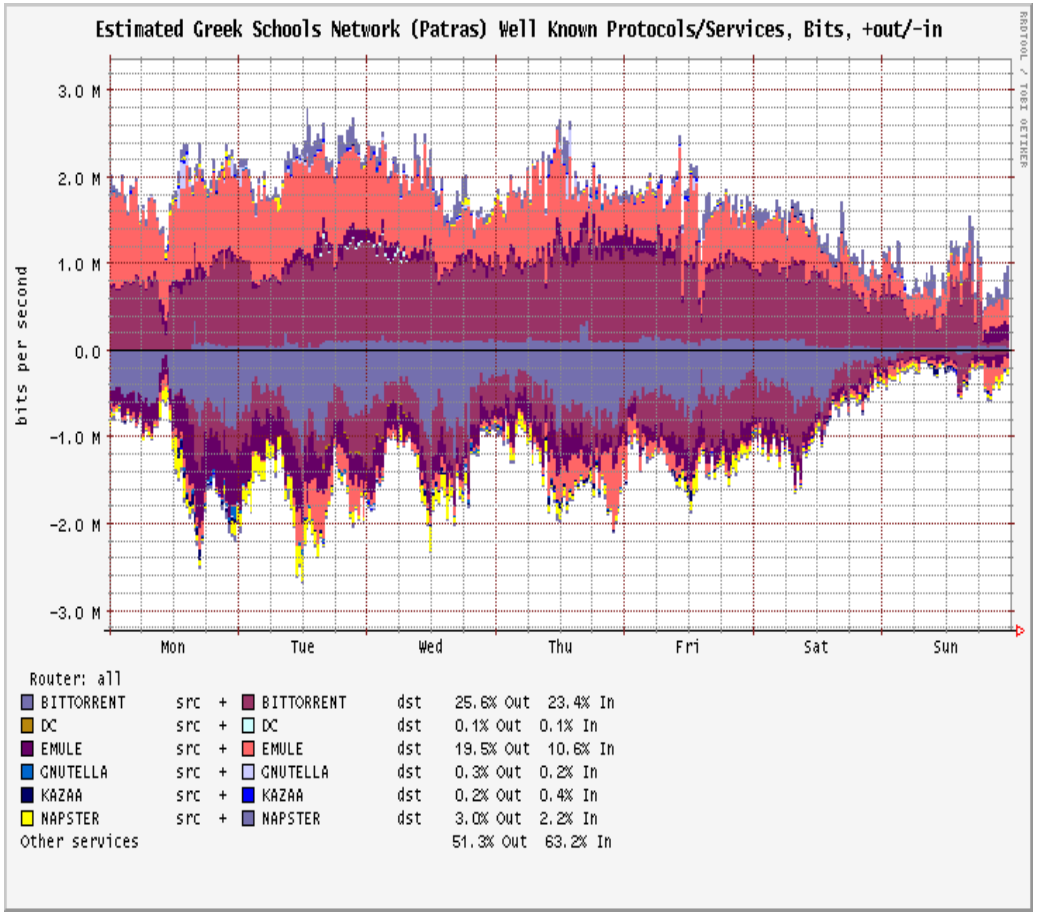
- Network Architecture
- Measurement Methodology
- Traffic Statistics
  - Service Analysis
  - Protocol Analysis
  - Flow Analysis
  - Packet Size Analysis
- Traffic locality
- Peer-to-Peer Services
- Conclusions

Peer-to-Peer Services

Protocol	outgoing traffic			incoming traffic		
	bits %	packets %	flows %	bits %	packets %	flows %
BitTorrent	25,6	17,9	5,9	23,3	18,7	7
eMule	19,5	16,1	12,8	10,6	14,3	14,6
Napster	3	2,5	0,3	2,2	2,3	0,4
Gnutella	0,3	0,3	0,2	0,2	0,3	0,2
Kazaa	0,2	0,2	0,1	0,4	0,2	0,1
Direct Connect	0,1	0	0	0,1	0	0
Total	48,7	37	19,3	36,8	35,8	22,3

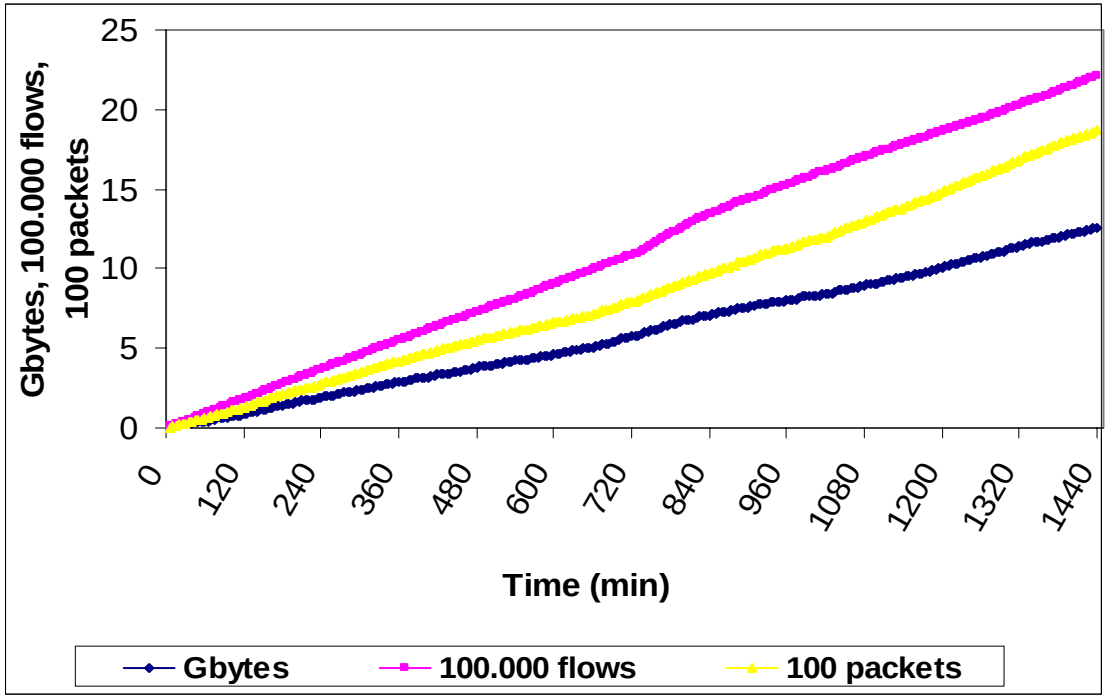
- Very Difficult to identify P2P traffic
  - The 3rd generation P2P systems use arbitrary ports for the P2P connections
- Still 25% of the traffic is unknown
- 32,3% - 48,7% of the outgoing and 14% - 39% of the incoming bytes are caused by P2P services

Peer-to-Peer Services



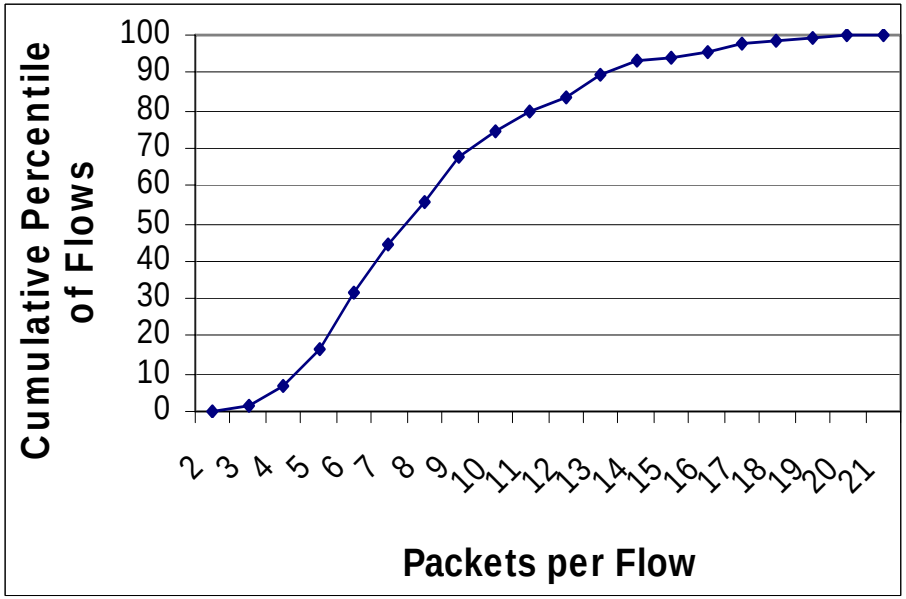
- P2P services are active 24 hours per day
  - + they do not follow the traffic pattern of the overall traffic
- Emule and BitTorrent were the two most prevalent protocols.
  - After 19/12/2004 the use of BitTorrent was reduced significantly because of the shut down of Suprnova.org

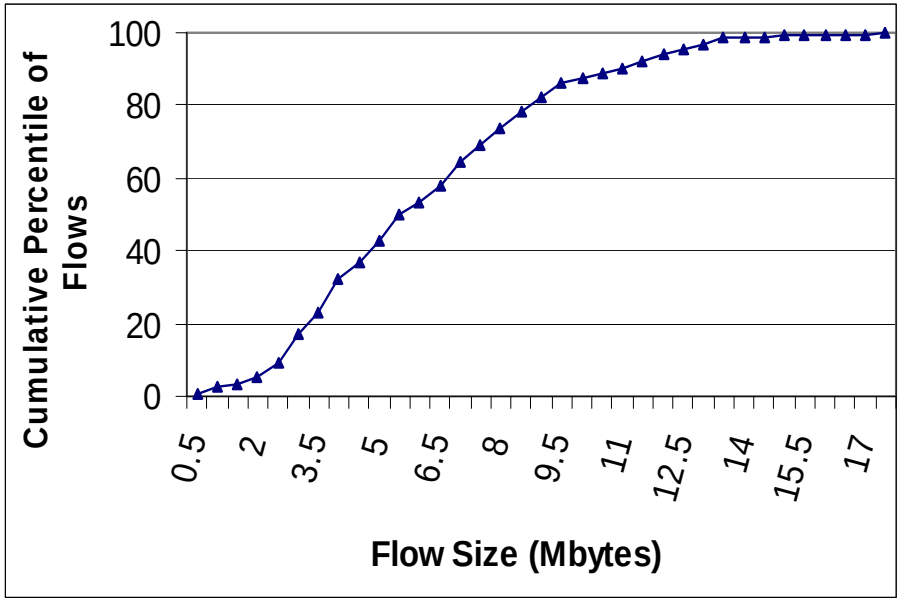
Peer-to-Peer Services



- The arrival rate remains relatively constant throughout the day.
- The same pattern on a weekly interval
  - A constant rate during the weekdays and a different rate (but constant again) during the weekends.

Peer-to-Peer Services





- The majority of P2P flows contain a relatively small number of packets.
- The average size of a P2P flow was 9 packets.
- P2P applications belong to the bulk data transfer-style applications.
- The mean P2P flow size is 6.1 Mbytes which is much bigger than the mean flow size of web traffic and other bulk data transfer services.

Overview

- Network Architecture
- Measurement Methodology
- Traffic Statistics
  - Service Analysis
  - Protocol Analysis
  - Flow Analysis
  - Packet Size Analysis
- Traffic locality
- Peer-to-Peer Services
- Conclusions

Summary – Future Work

- The traffic has daily and weekly periodic components as well as a long-term trend
  - Non-stationary model
- The traffic is increasing in time
  - In 4.5 months 100% increment of traffic rate during the peak hours
- TCP by far dominates the network traffic.
- HTTP and P2P services are the most frequently used applications
  - have to be taken into account in a future network extension
  - Tools like FlowMonitor have to be implemented
- Strong traffic locality phenomenon. 1% of the sources correspond to 95% of the outgoing bytes.
- The majority of the flows last a few seconds and carry few packets
- Predominance of small packets
- Apply realistic models that captures all the trends of the traffic.

ACADEMIC RESEARCH COMPUTER TECHNOLOGY INSTITUTE

CTI

Analyzing Traffic across the Greek School Network

Thank you!

LANMAN 2005, 14th IEEE Workshop on Local and Metropolitan Area Networks,
18-21 September 2005, Chania, Crete, Greece