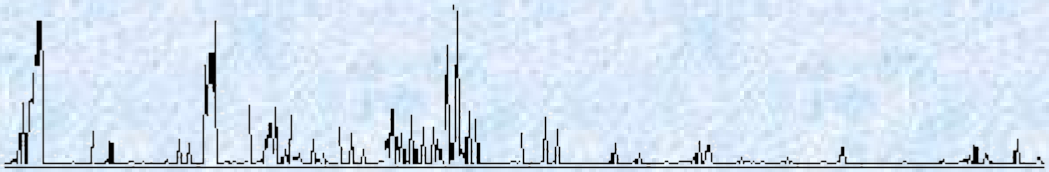
Traffic Flow Confidentiality in IPsec: Protocol and Implementation

Eavesdropping

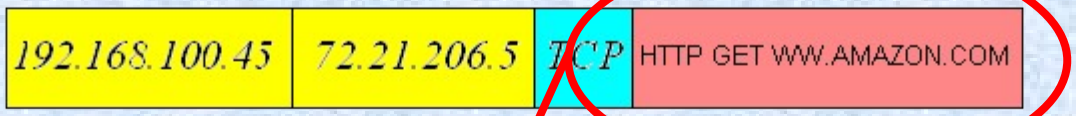
client
192.168.100.45



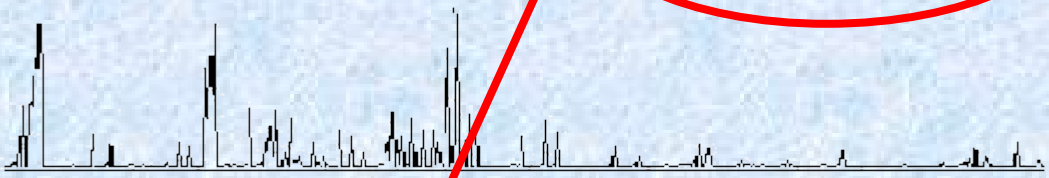
www.amazon.com
72.21.206.5

EAVESDROPPER

Eavesdropping



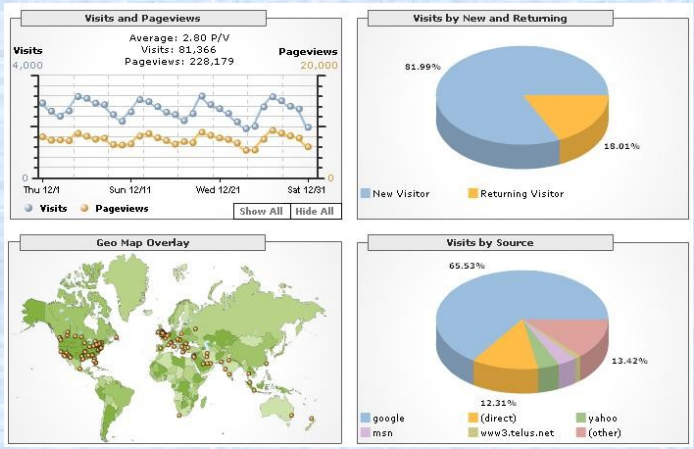
client
192.168.100.45



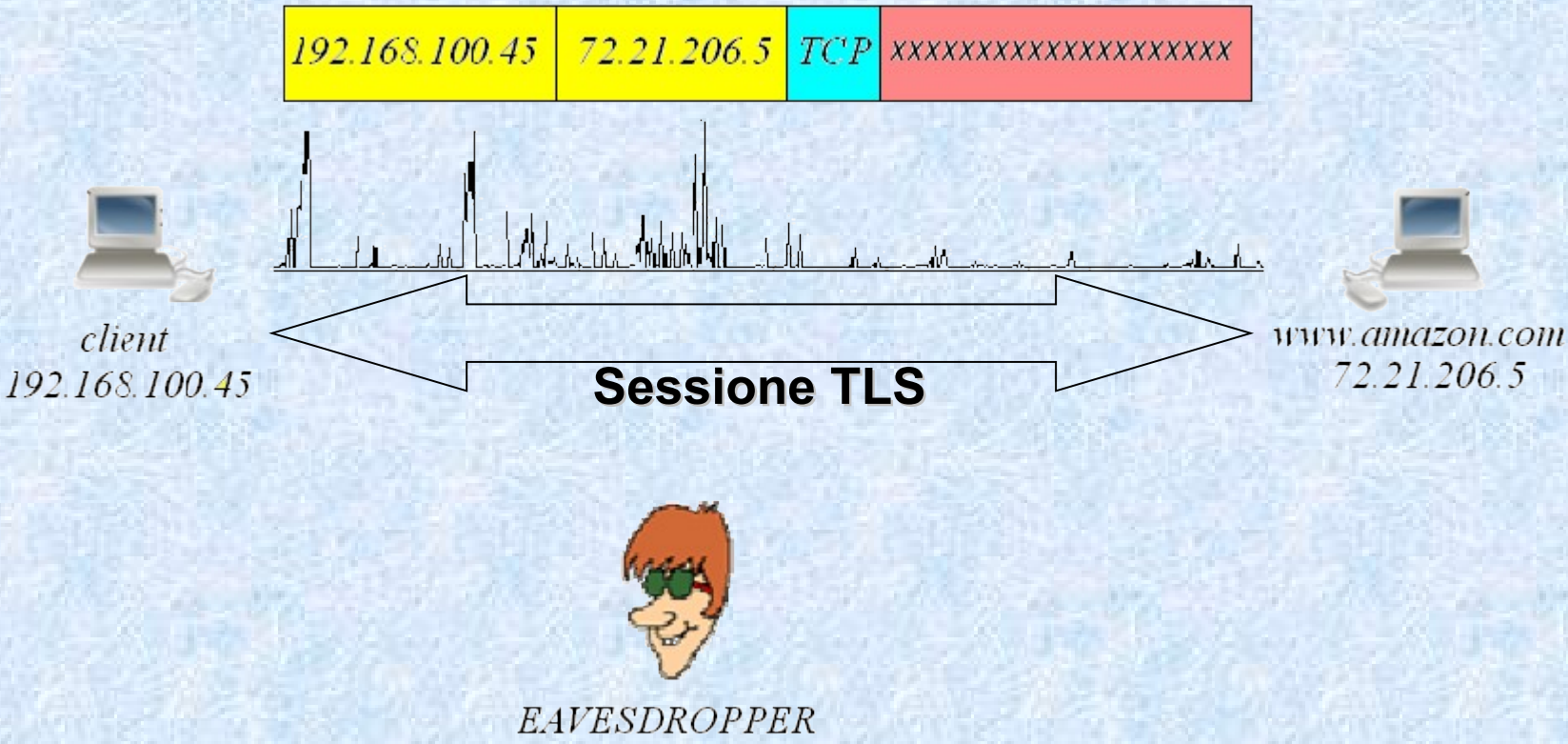
www.amazon.com
72.21.206.5

EAVESDROPPER

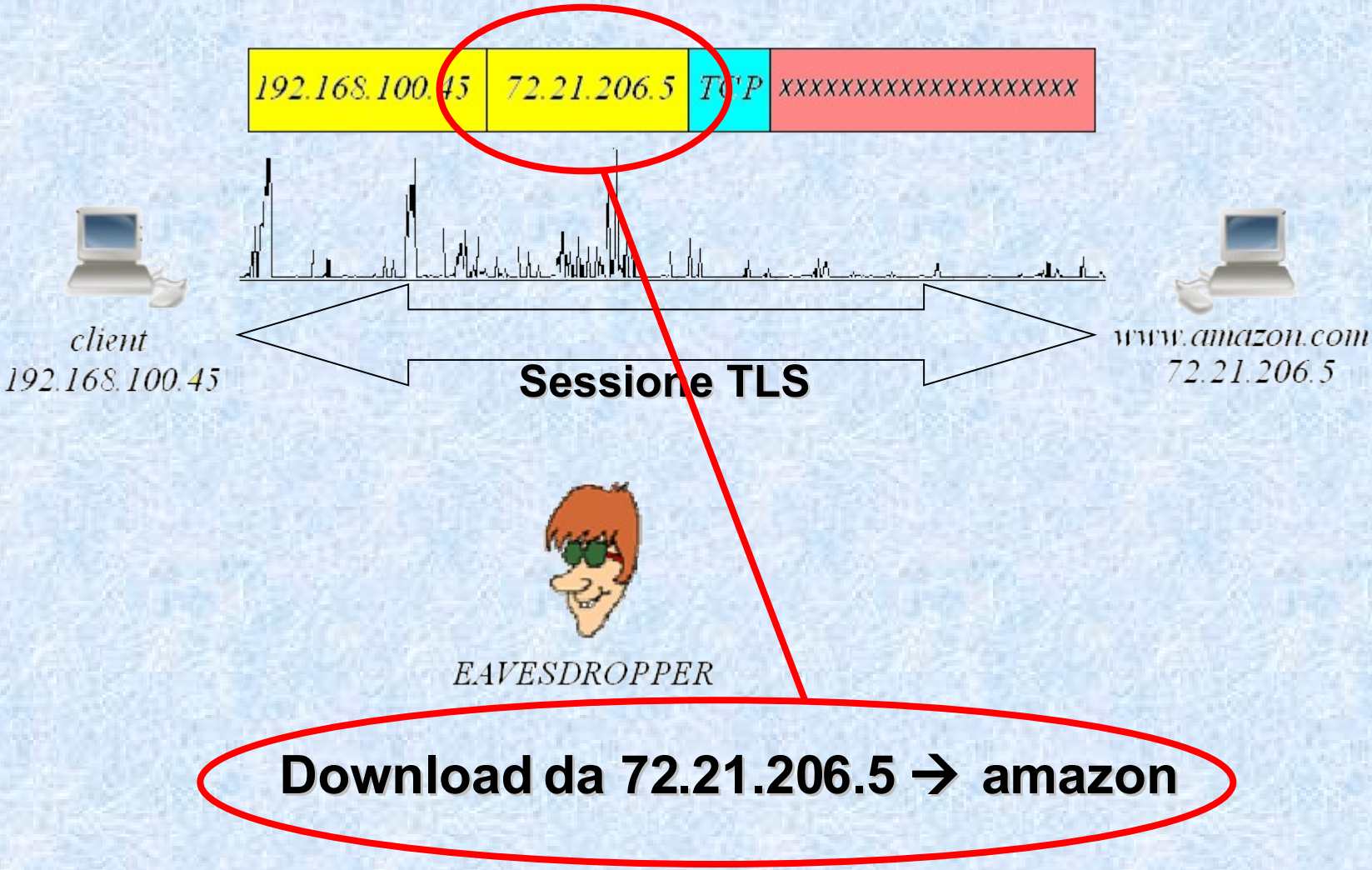
Download da amazon



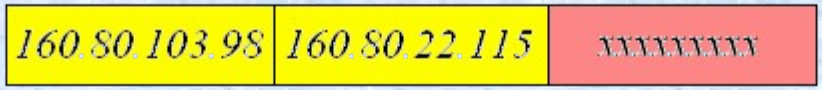
Linkability

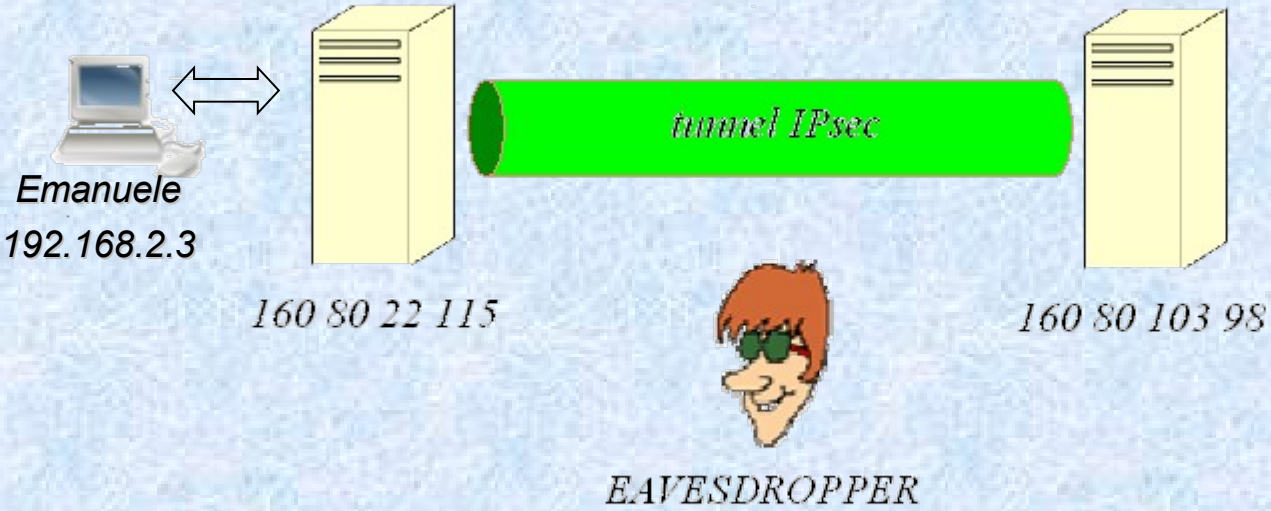


Linkability

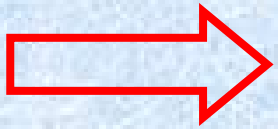


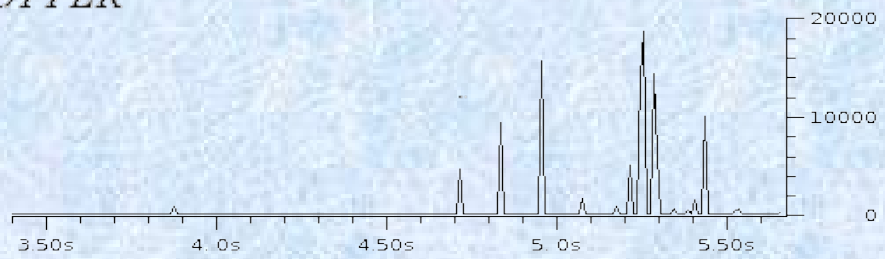
Malicious Traffic Analysis





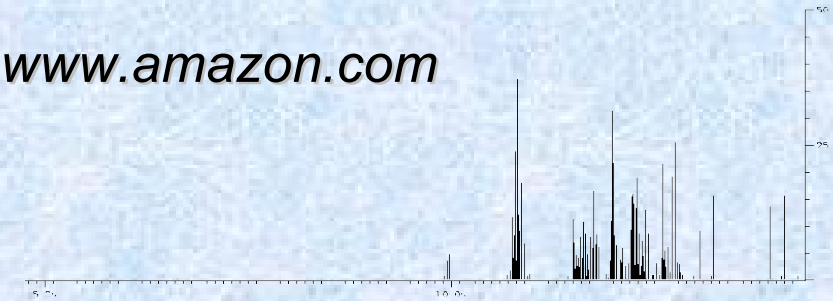
- Length
- Arrival time
- Packets direction



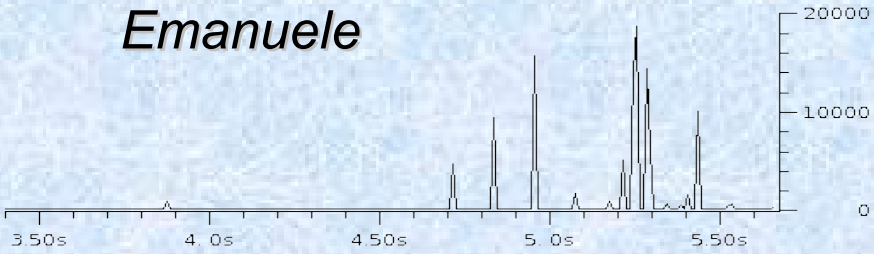


Source-Destination Link

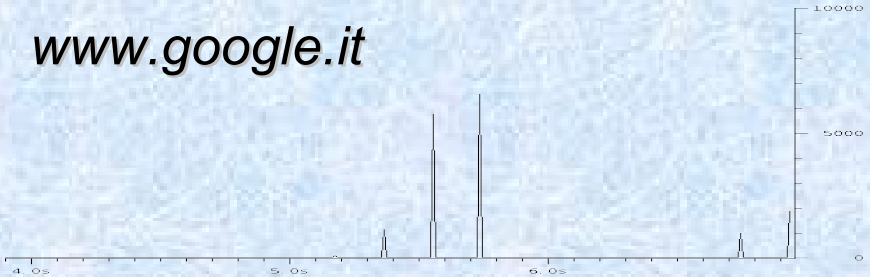
www.amazon.com



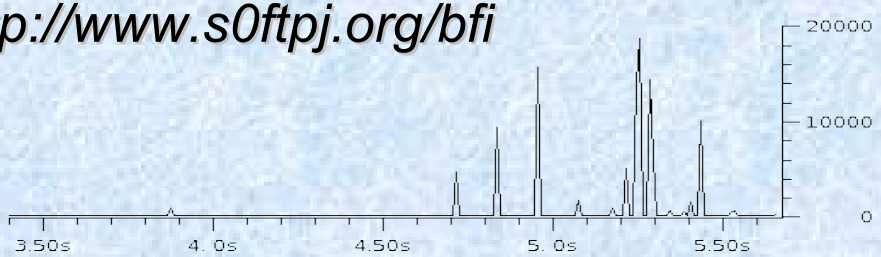
Emanuele



www.google.it



http://www.s0ftpj.org/bfi

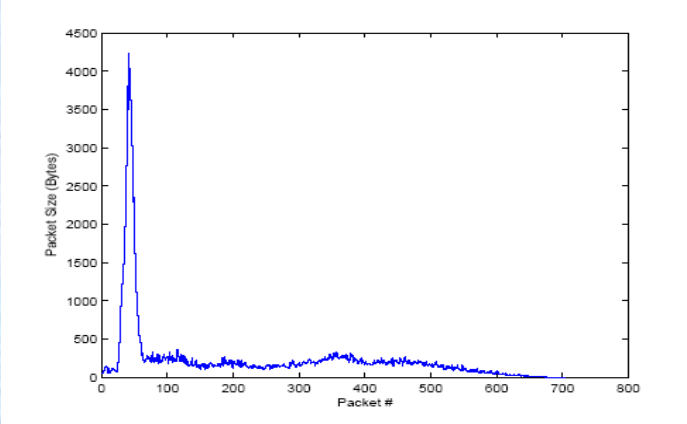


Traffic Analysis misuse

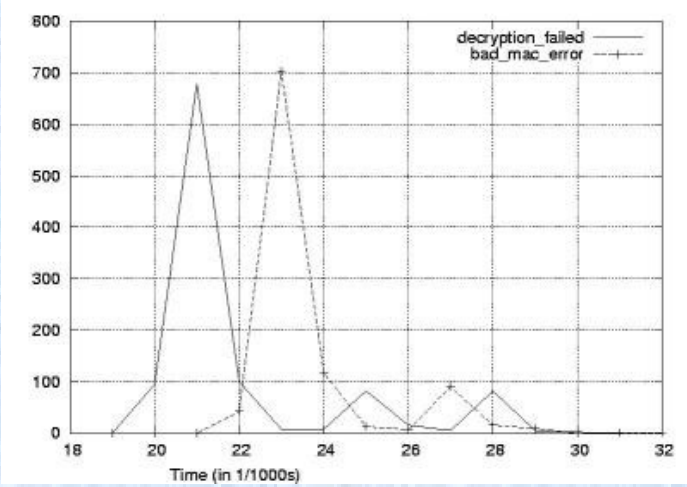
- Traditional attack support
  - Password recovery
  - Information recovery
- Attacks against Privacy
  - Web site fingerprinting
  - Protocol fingerprinting
- Attacks against the anonymization network
  - Correlation attack
  - Latency attack

User information recovery

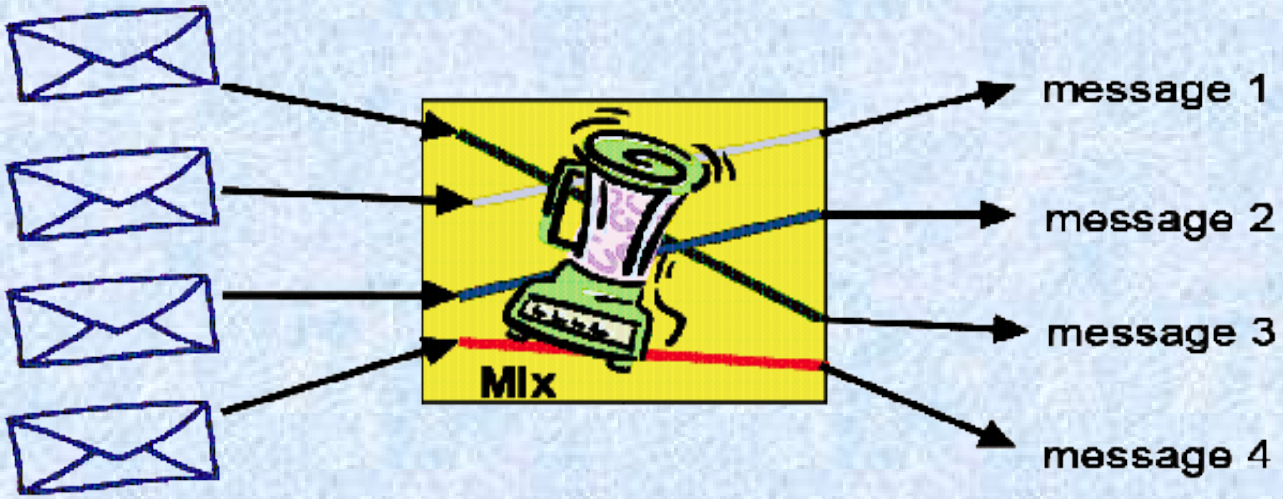
- Website fingerprinting
  - E.g. sample size profile for www.amazon.com
  - Bissias, Liberatore, Levine “Privacy Vulnerabilities in Encrypted HTTP Streams”



- Password recovery
  - Canvel, Hiltgen, Vaudenay, Vuagnoux, “timing-based attack to Intercept passwords in a SSL/TLS Channel”
    - Different log-in error are characterized by different server’s answer times
    - http://www.brice.info/crypto



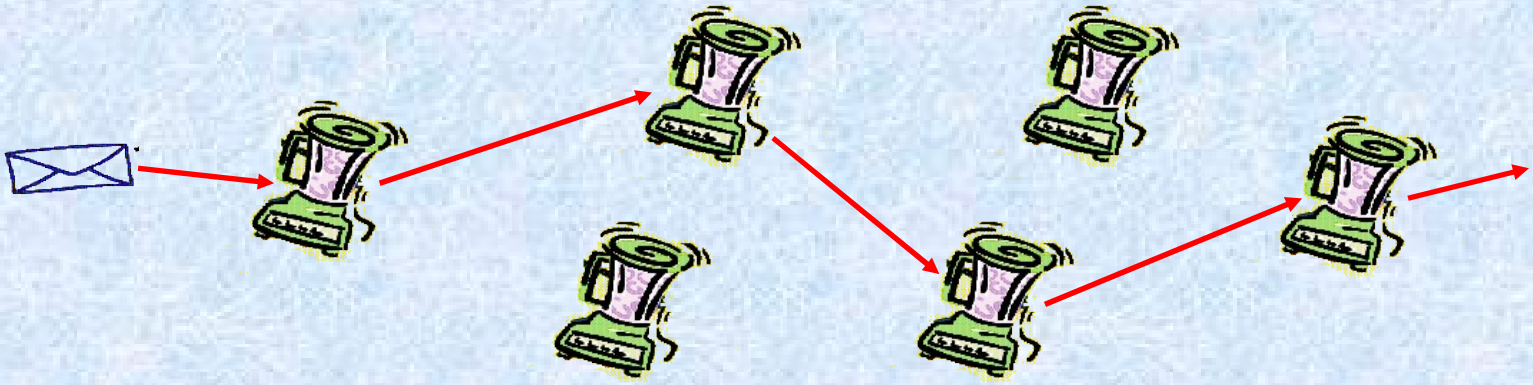
MixNet basic ideas



Messages:

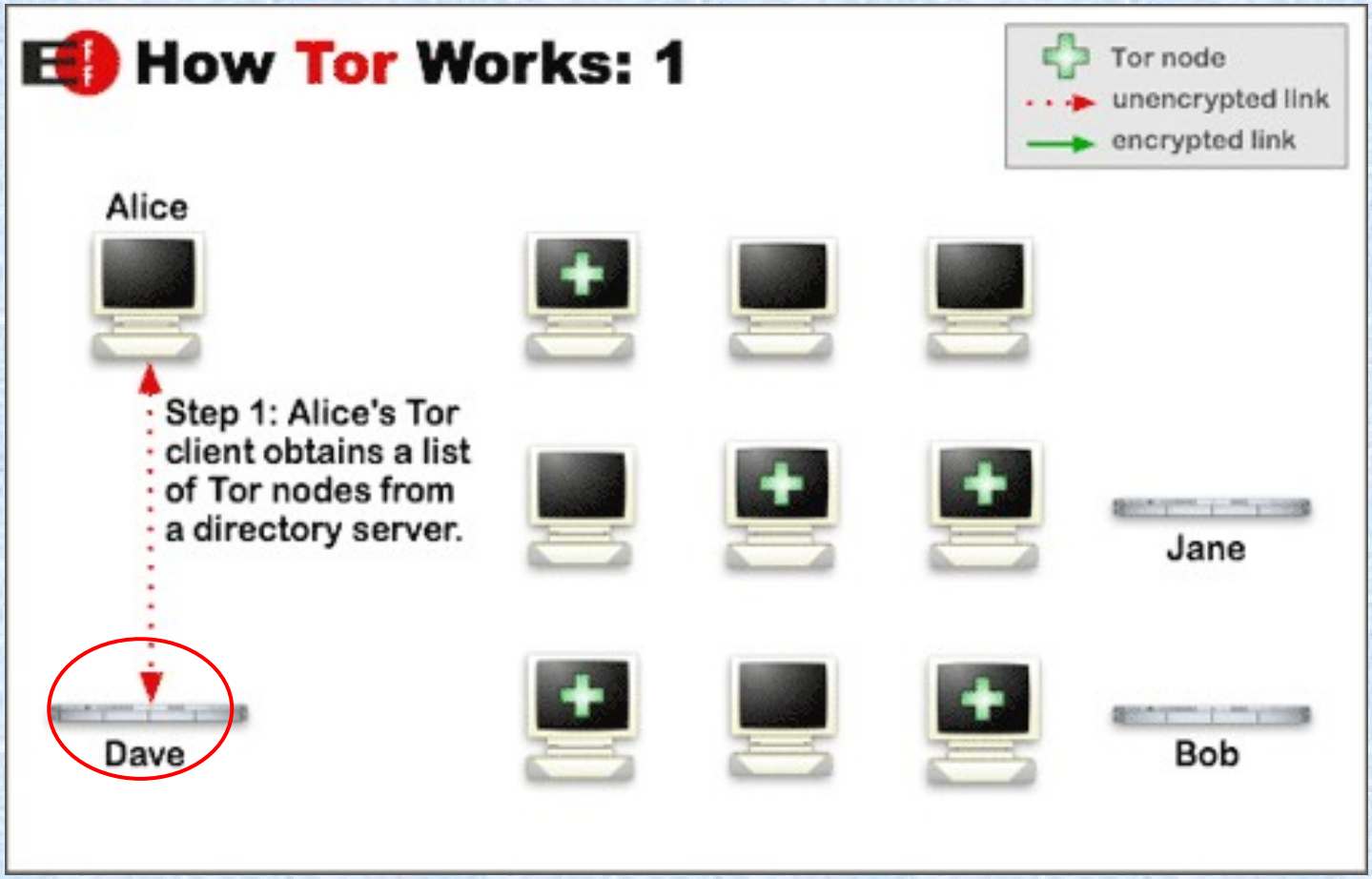
- wrapped in fix length packs
- grouped and sent in lexicographical order
- in/out correspondence hidden by mix

- "Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms," D. Chaum
- Employs a “network” of mixes to avoid the need of a single trusted one



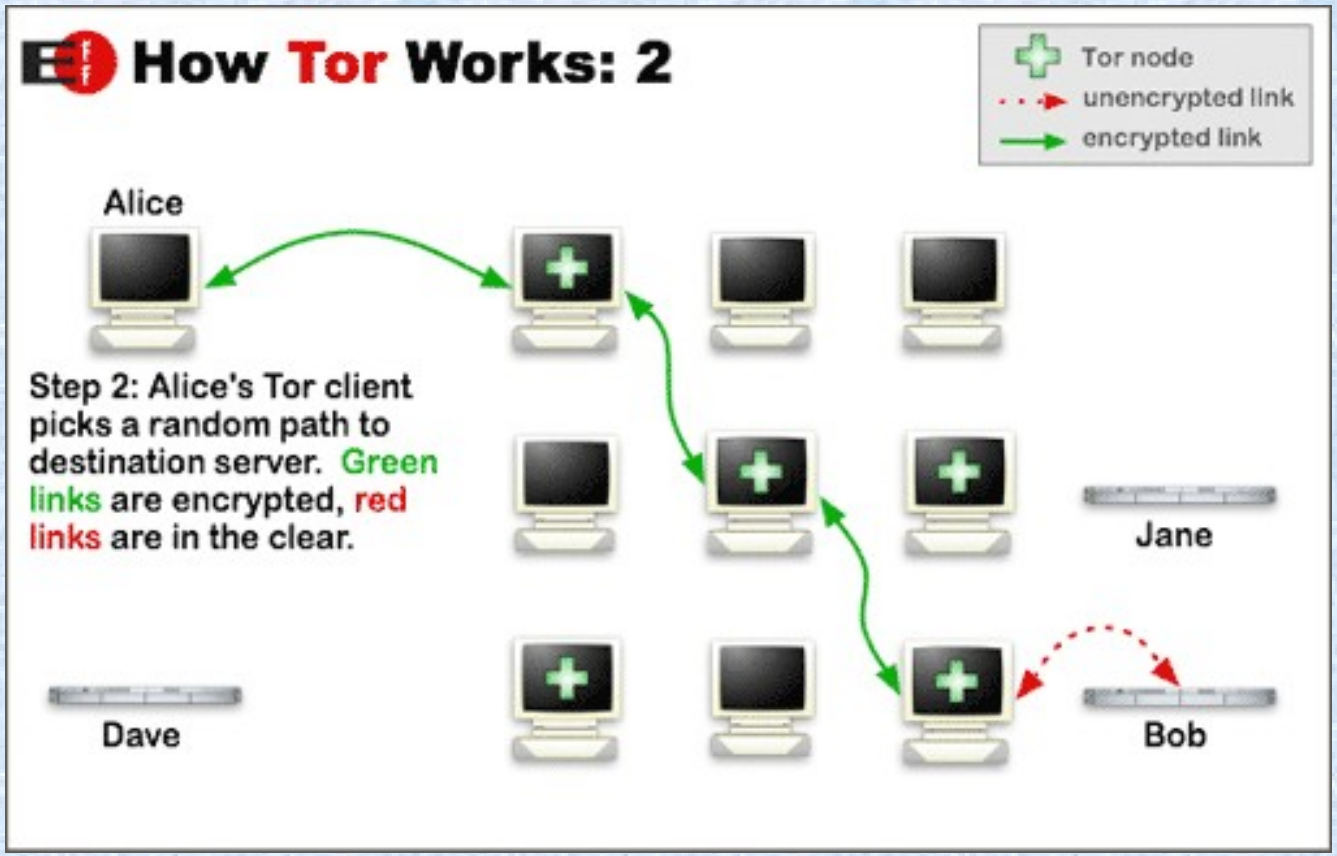
Tor 1/3

- Directory server : maintain dawned informations about network topologyand nodes state



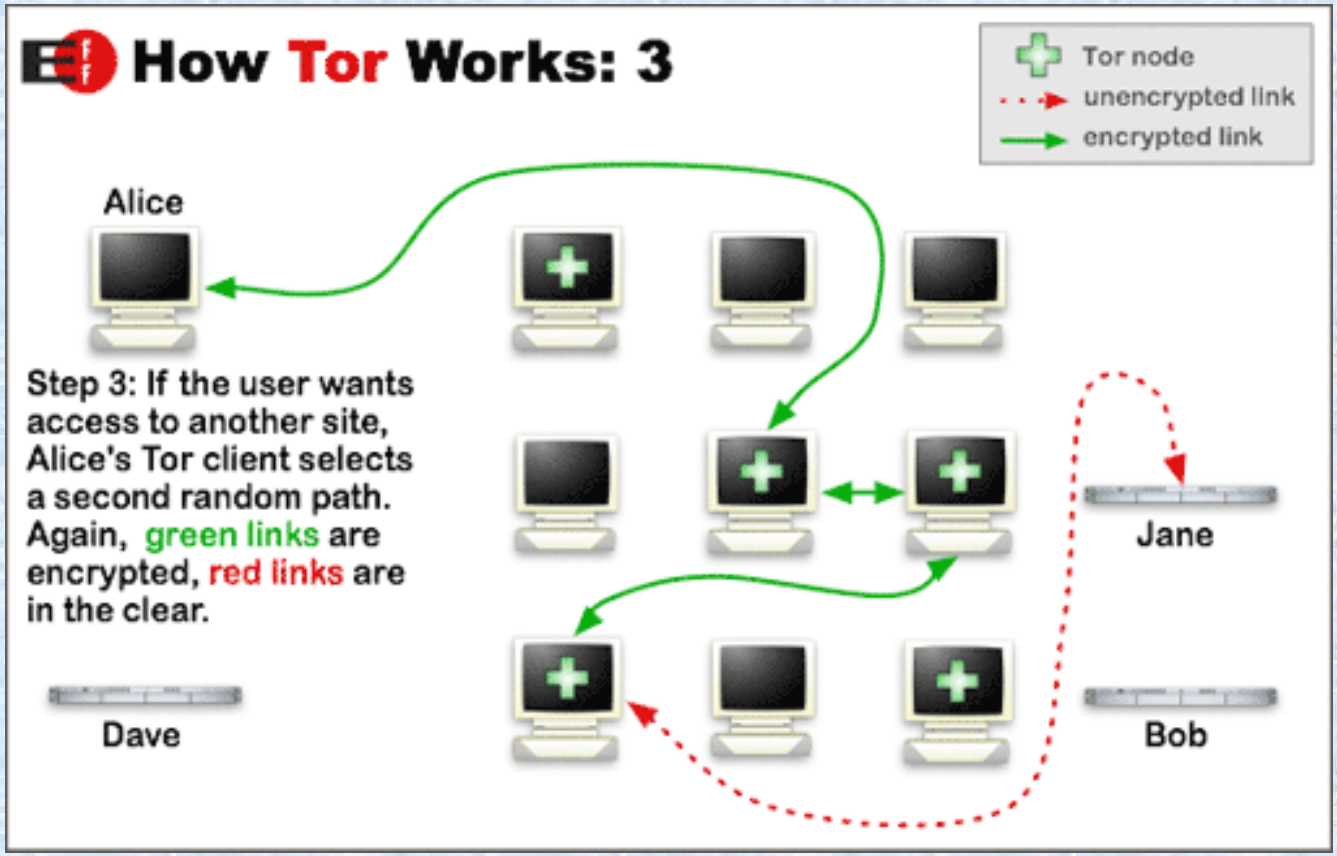
Tor 2/3

- Every Onion Router (OR) does't know the complete comunication path



Tor 3/3

- Following comunications use different path to avoid correlation with old data exchanges



Anonymization network limit

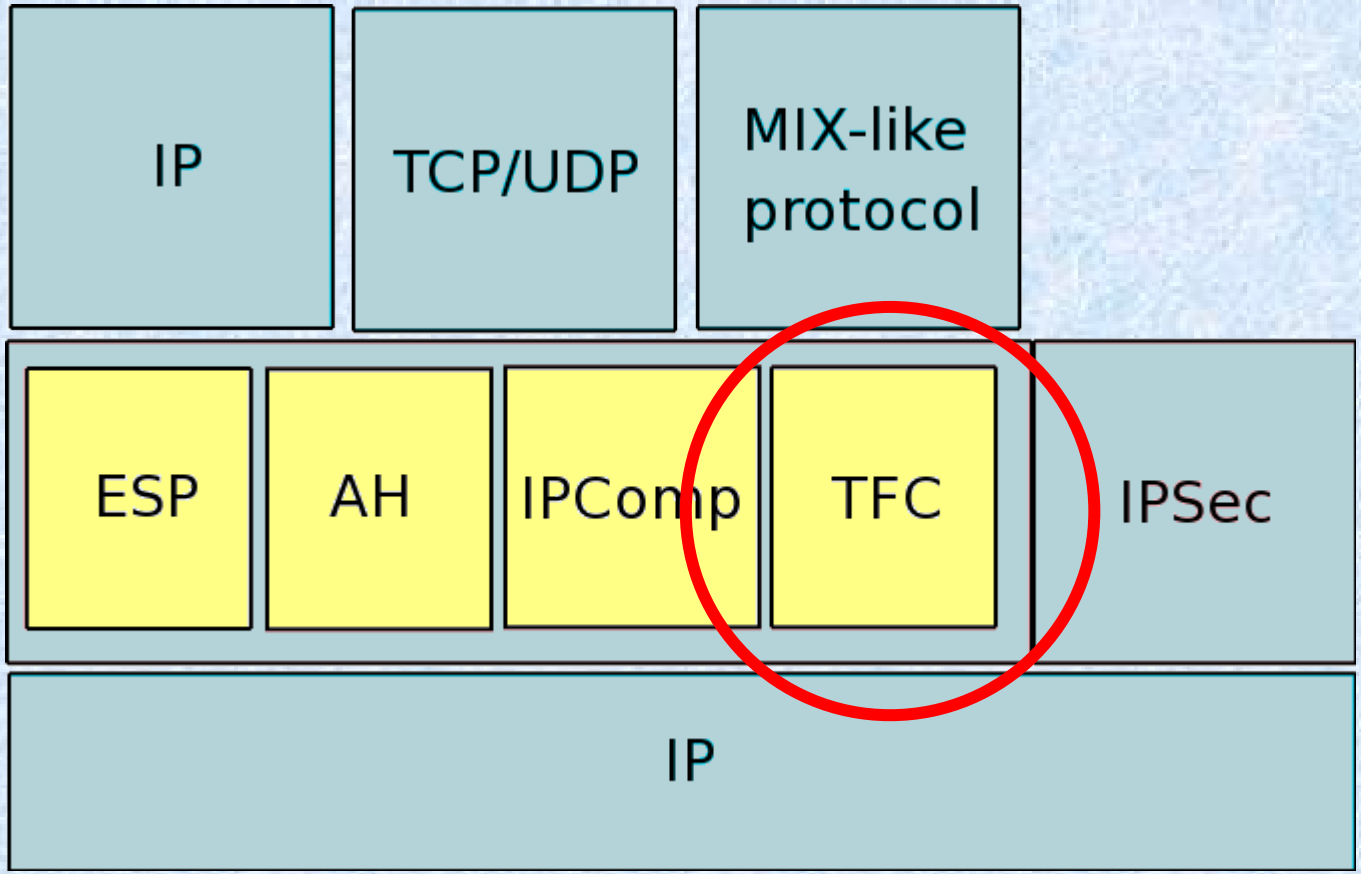
- Monolitic design and implementation:
  - Every anonymization network provides all togheter:
    - Channel protection mechanisms
    - Source routing mechanism
    - Topology discovery mechanism
    - Signaling protocol
  - Support for specific kind of traffic
    - Real time traffic
      - Tor
    - Mail
      - Mixminion
      - Mixmaster
    - P2P
      - Tarzan
      - MorphMix

Goals

Our goals is to provide a tool:

- implementing the basic mechanisms to prevent statistical traffic analysis attacks
  - Dummy traffic
  - Packets padding
  - Traffic re-shaping
- Flexible
- Reconfigurable
- Reprogrammable
- Based on common standard
  - IPsec
- Providing a underlying layer for the Anonymous Routing Networks
  - Supporting different kind of traffic

Traffic Flow Confidentiality Protocol



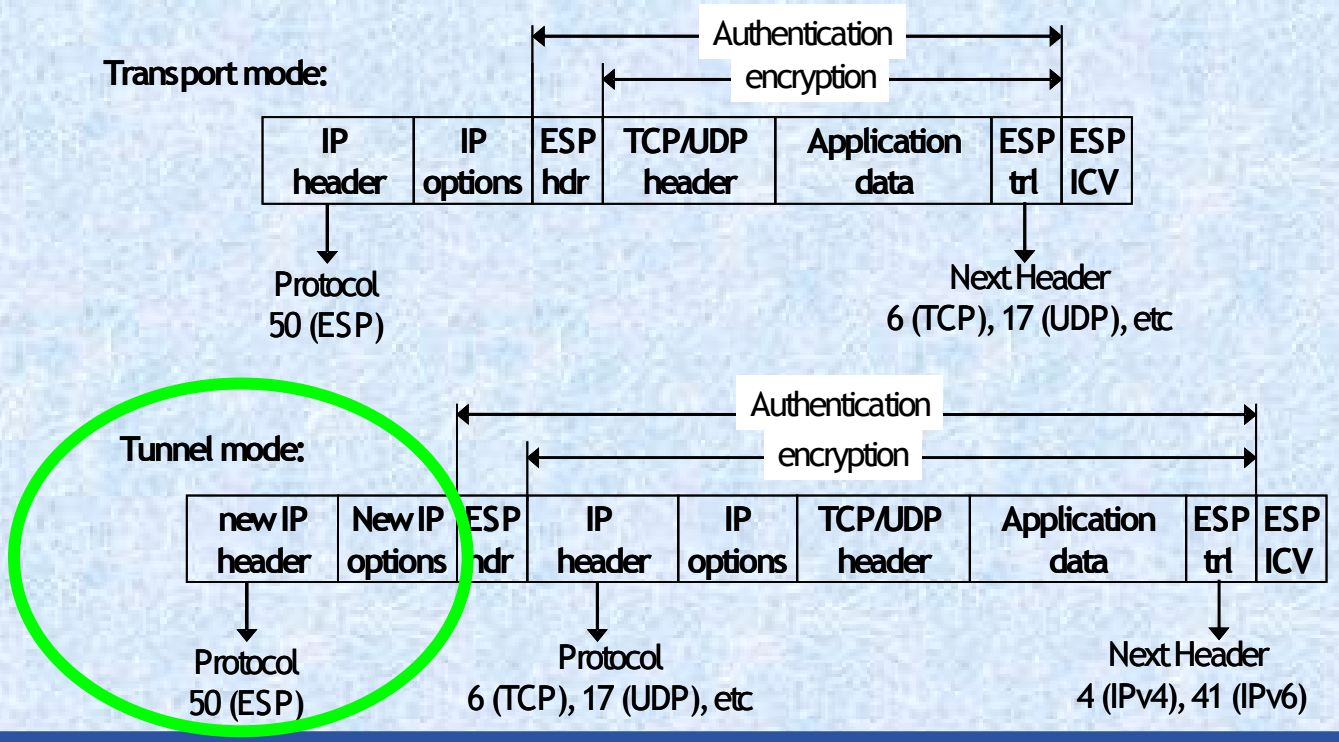
TFC, like ESP and AH, can be managed exploiting the instruments Offered by IPsec (SA, SAD, SPD, …)

Review of IPsec basics

- Universal → protection at IP layer, automatically protects all applications
- Protects IP header and payload
- Prevents replay attack, Denial of Service
- Standard (RFC4301-4306)
  - The only known protocol to include limited Traffic Flow Confidentiality mechanisms in its more recent specification
- Versatile – different options, tradeoff
  - Two modes of operation: Transport, Tunnel
  - Two protocols: AH(authentication), ESP (encryption + authentication)
  - IKE: optional, flexible key management protocol

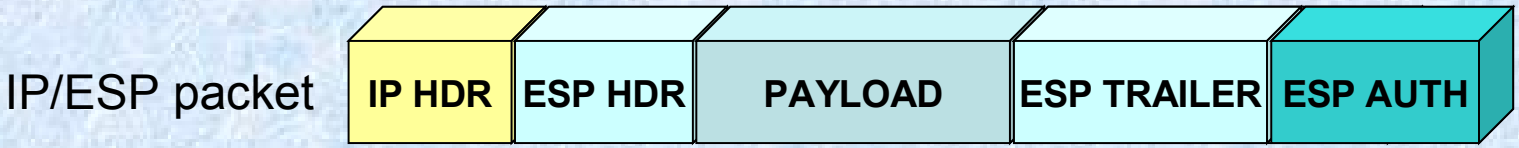
Encapsulating Security Payload (ESP)

- Must encrypt and/or authenticate in each packet
- Encryption occurs before authentication
- Authentication is applied to data in the IPsec header as well as the data contained as payload



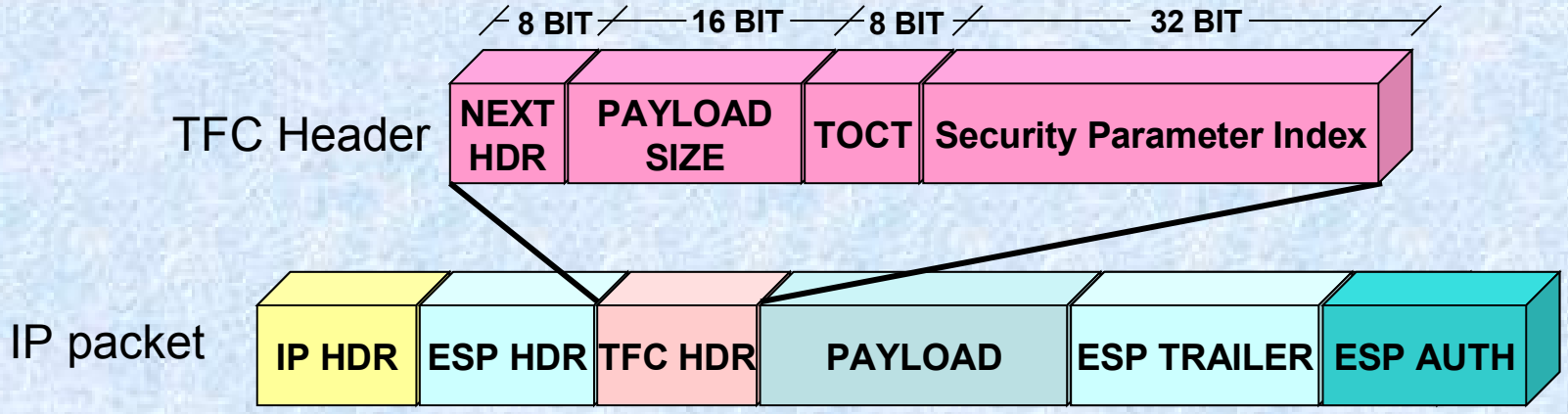
TFC Header

- TFC protocol header contains
  - Security Parameter Index (SPI)
  - Protocol transported
  - Size of the data
- The header is inserted between the ESP header and the payload
- The padding is added between the payload and the trailer ESP



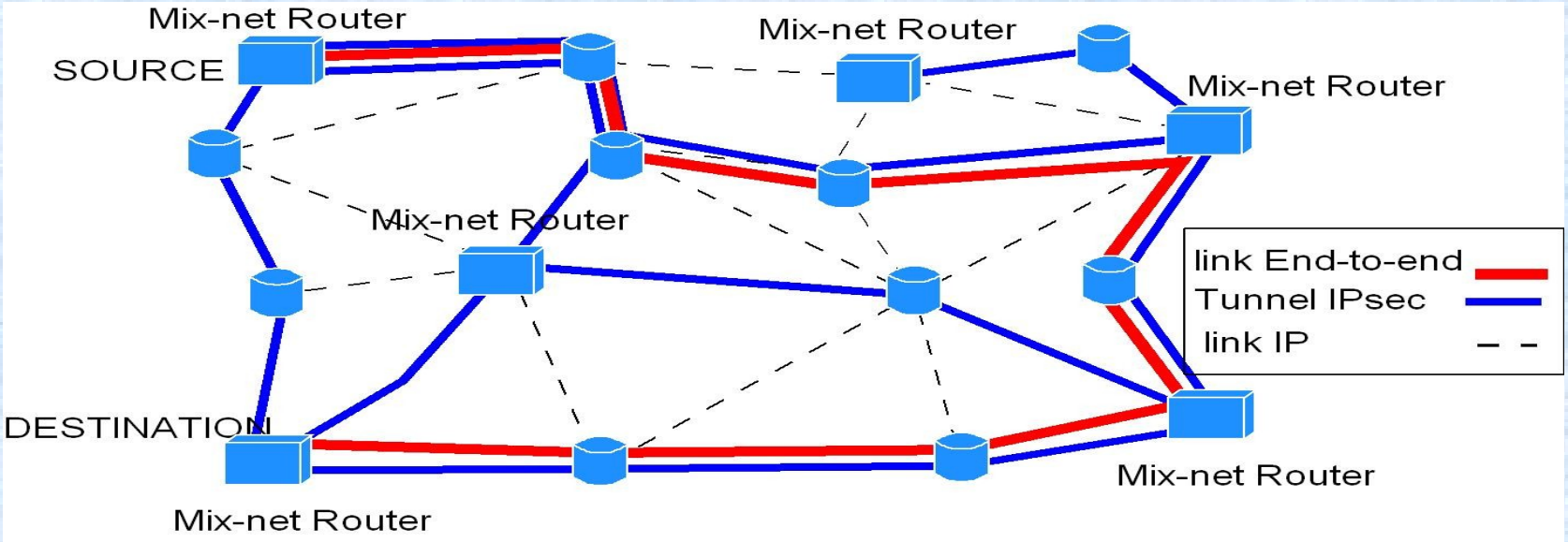
TFC Header

- TFC protocol header contains
  - Security Parameter Index (SPI)
  - Protocol transported
  - Size of the data
- The header is inserted between the ESP header and the payload
- The padding is added between the payload and the trailer ESP



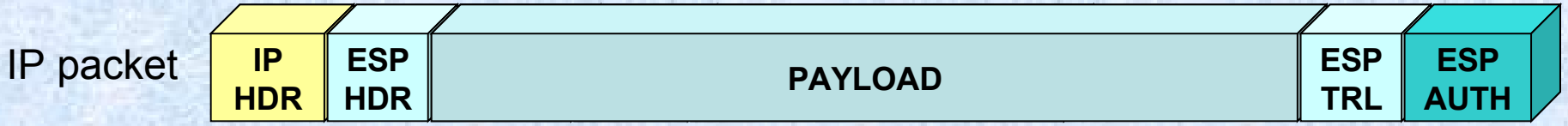
TOCT- Type of Confidentiality Treatment

- TOCT (Type of Confidentiality Treatment)
  - carry information about the type of treatment the packet may be subjected to
  - used in a multi-hop fashion, and especially for building IPsec-based Mix Networks.
- Still to evaluate information disclosed!!



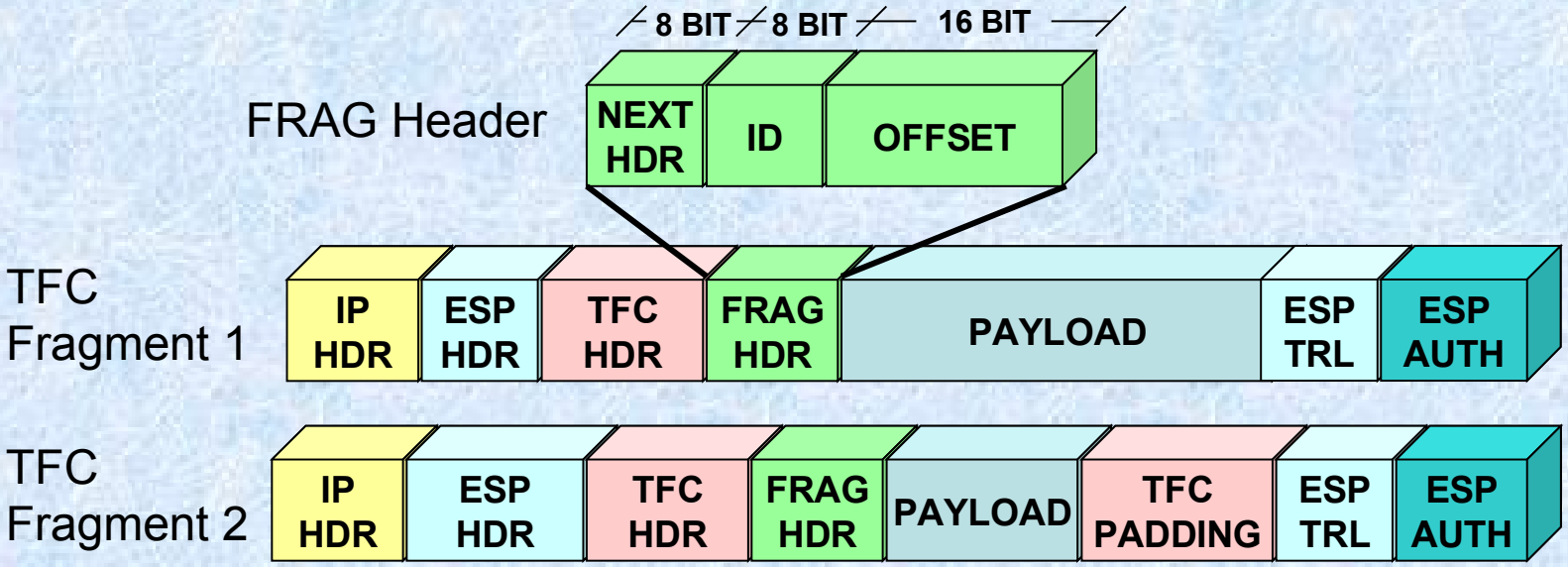
Packets fragmentation

- It has been necessary to add an extension header (FRAG)
- If needed the last fragment is padded

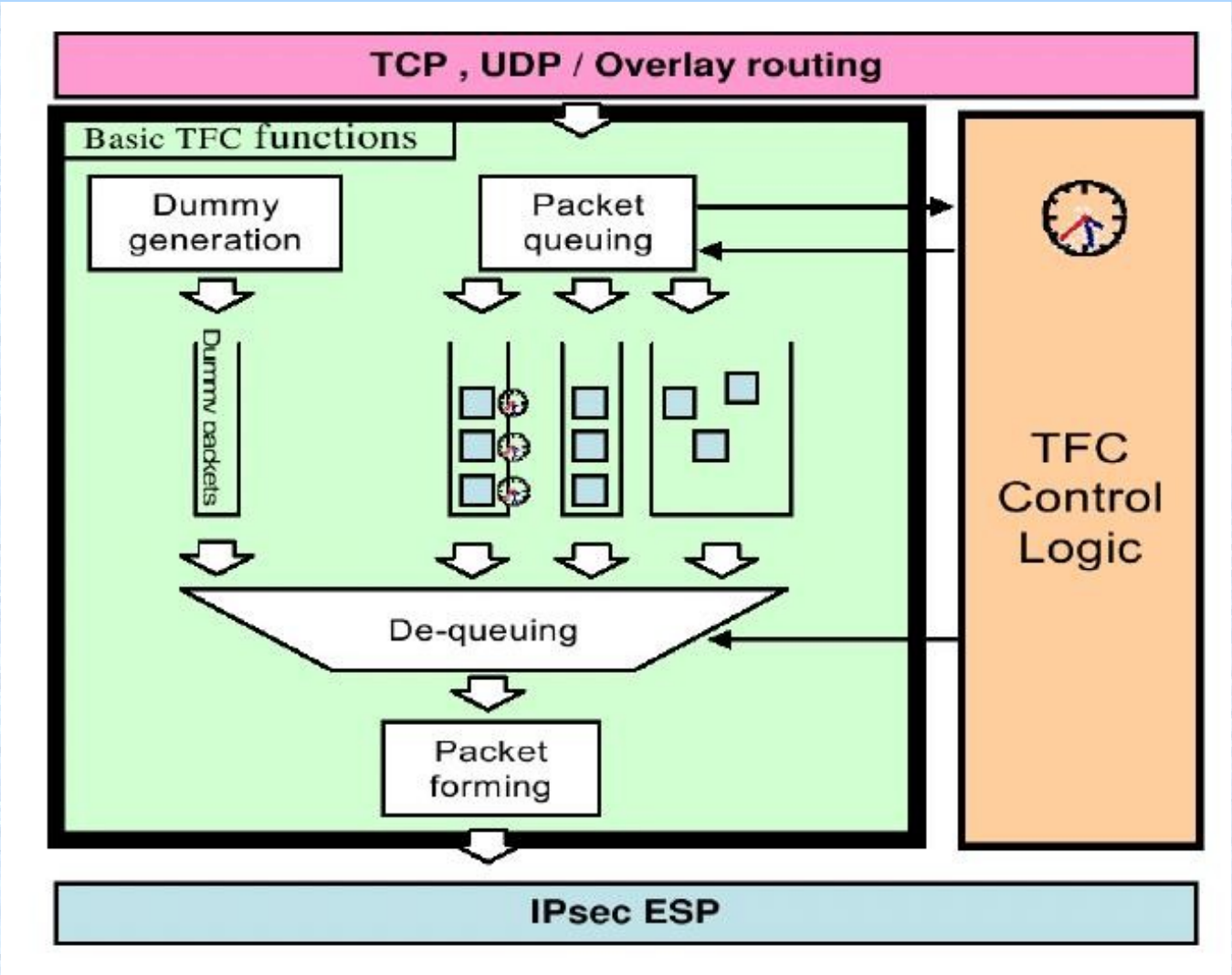


Packets fragmentation

- It has been necessary to add an extension header (FRAG)
- If needed the last fragment is padded

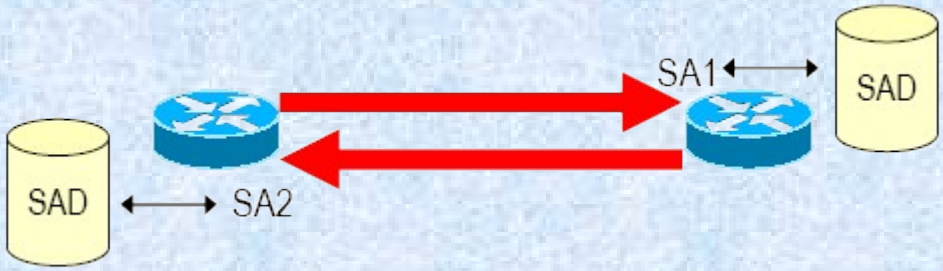


TFC architecture



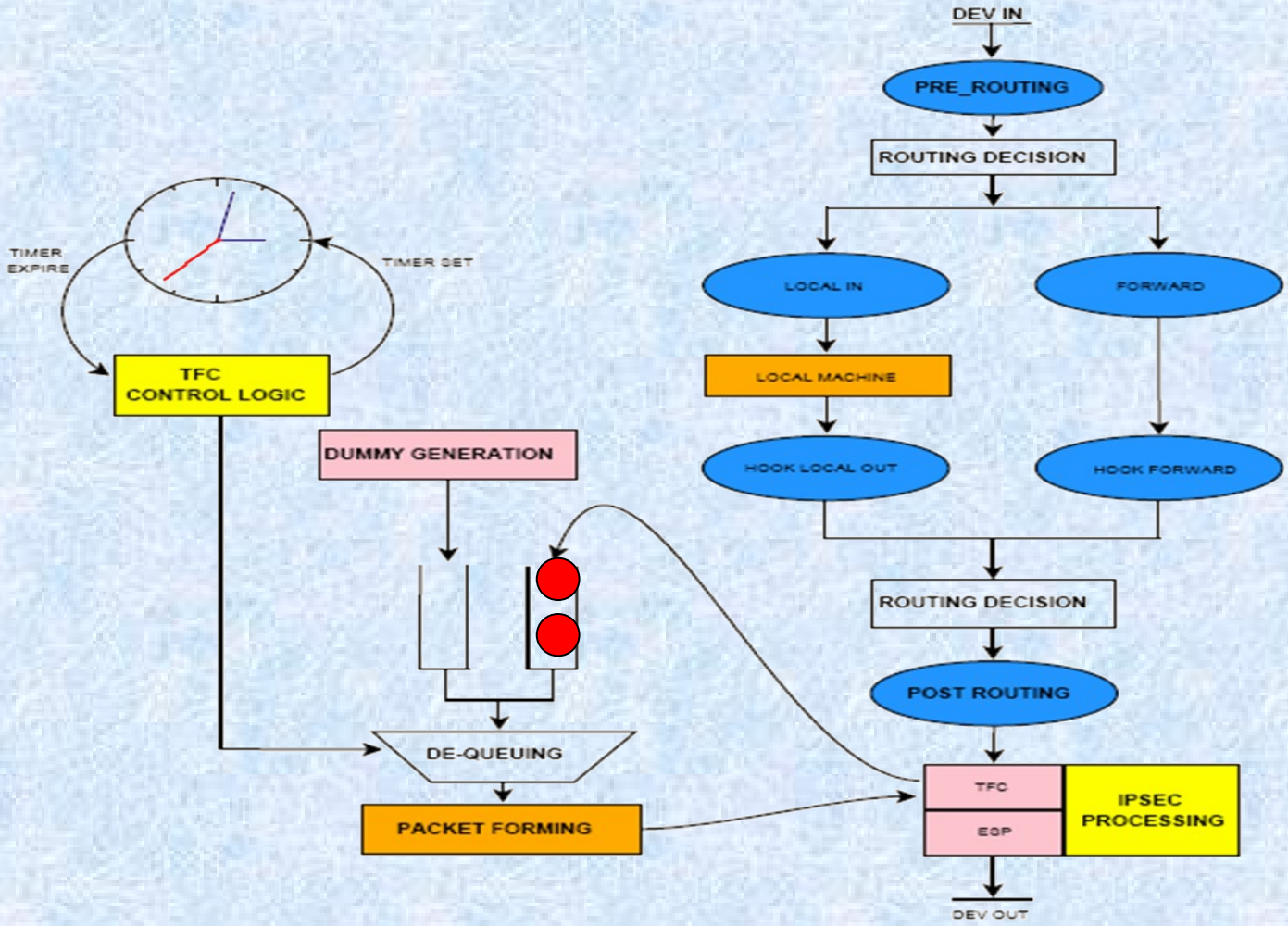
TFC SA parameters

- A User Space application allows to configure TFC SA parameters
  - Delay Algorithm
  - Dummy
  - Padding
  - Fragmentation
  - Packets Length
  - Bit Rate

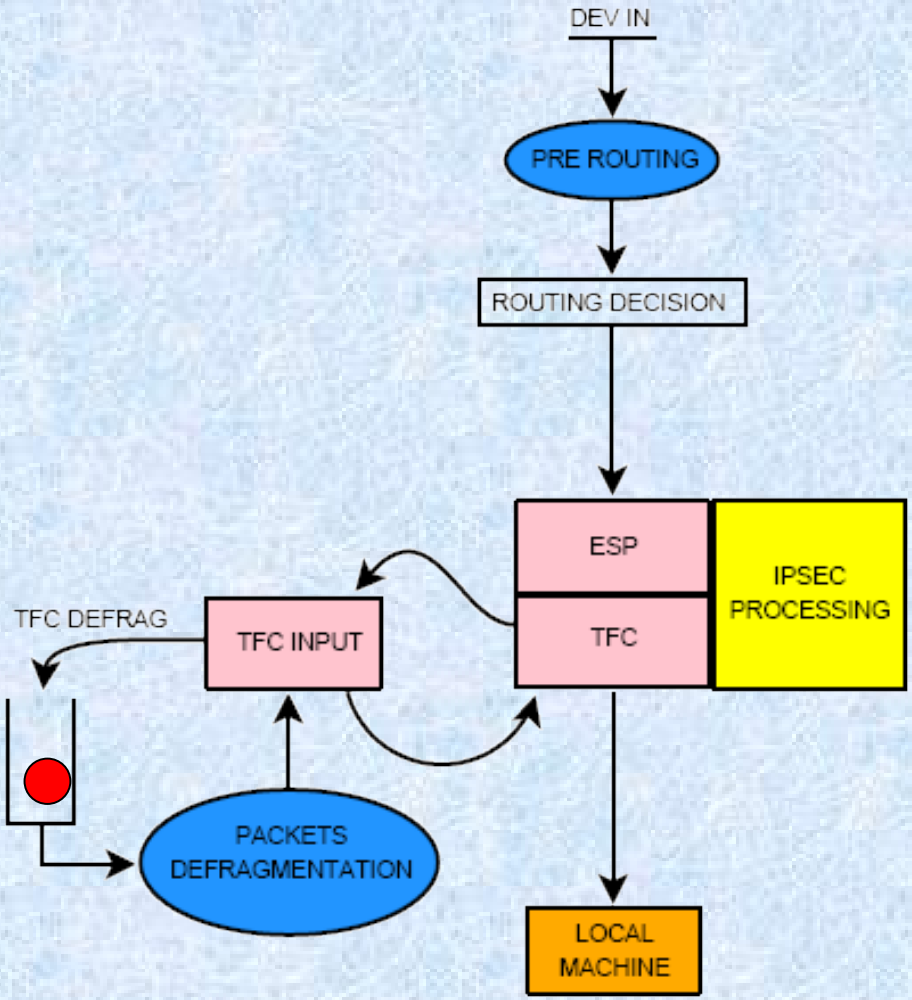


- Data are sent from the user space to the kernel space with a netlink socket

Packets Output Stack

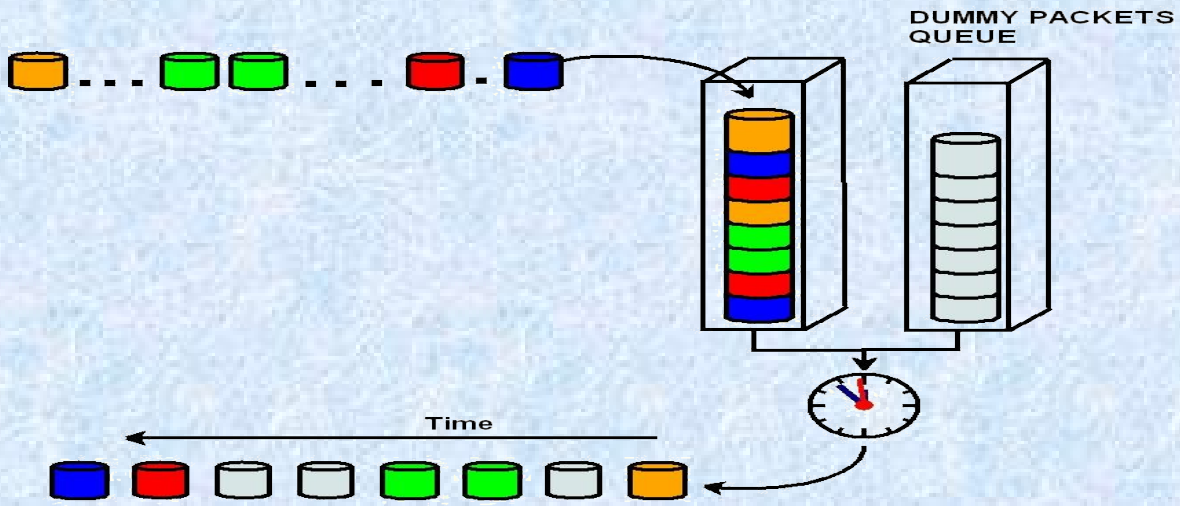


Packets Input Stack

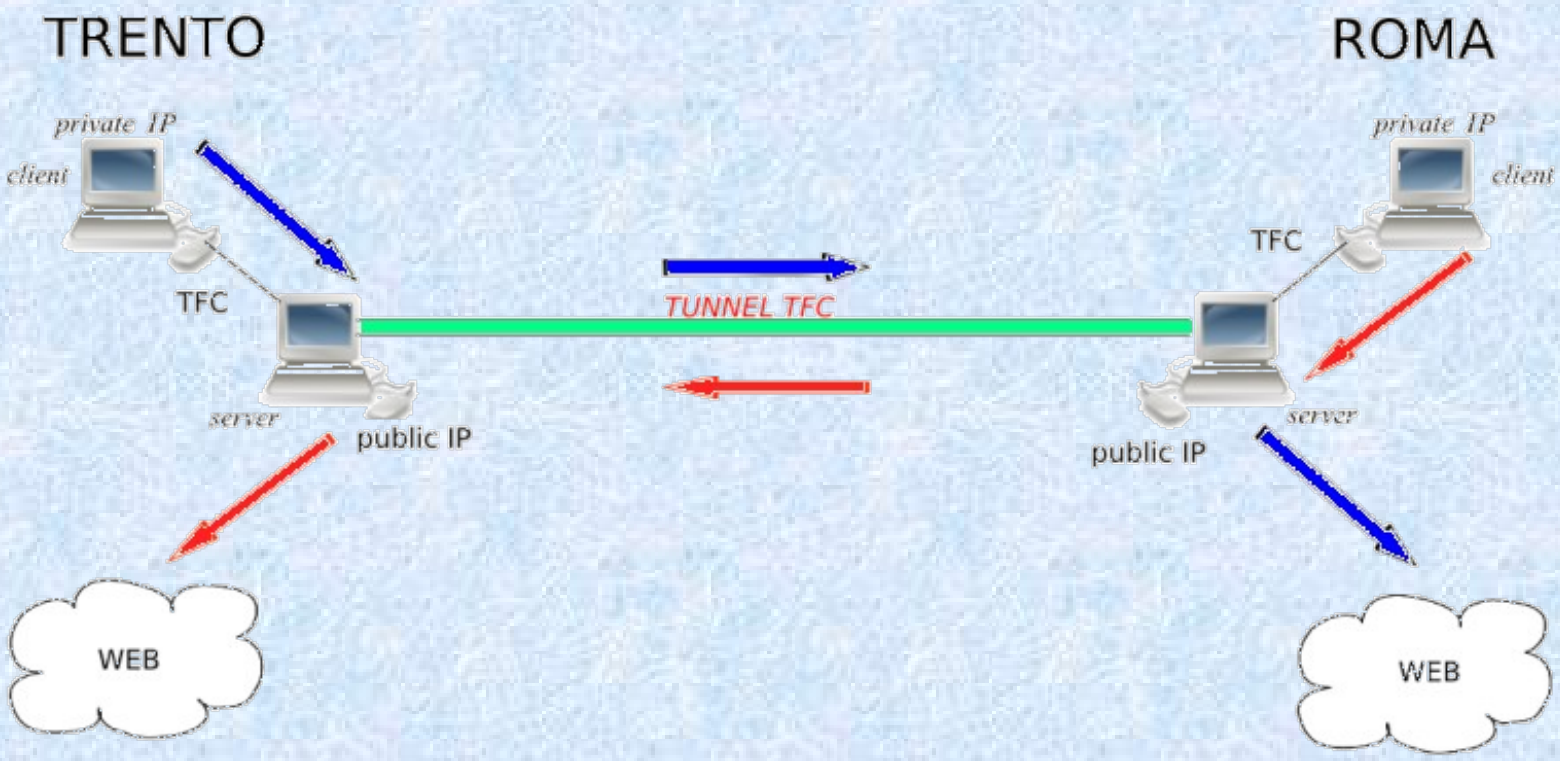


Timer and Dummy packets

- Timer associated with a TFC Security Association(SA). When the timer expires one or more packets are sent
- If the packets queue is empty one or more dummy are sent (IP protocol = 59)



Test su rete pubblica
Tunnel Roma - Trento

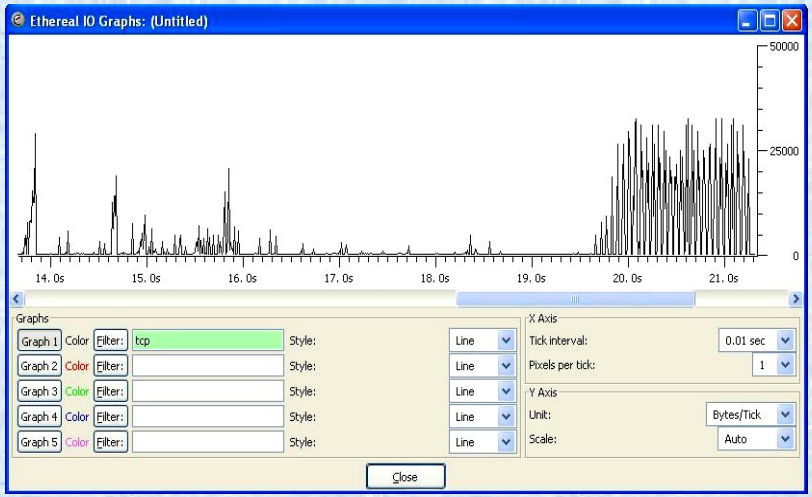


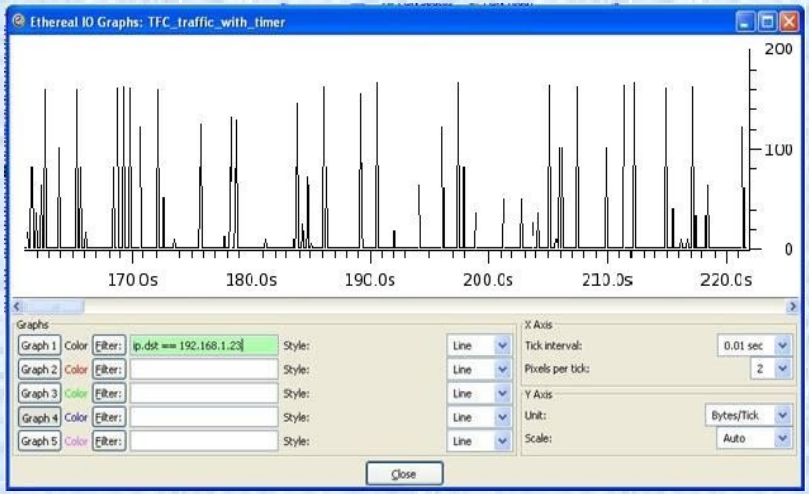
Control Logic

- The "control logic“ is the "intelligence" of the system
- It can combine the TFC basic mechanisms arbitrarily:
  - batching,
  - CBR (Continuous Bit rate),
  - random padding,
  - random delay algorithms
  - Queue congestion Reactive algorithm (still experimental)
- Simple methods (fixed or random packet clocking), may be easily replaced by more complex algorithms
  - Able to take into account the status of the queues and/or the congestion level
- The effectiveness of such adaptive approaches in terms of performance/privacy gains and trade-offs is still to be assessed

TFC flows sample

- We tested the TFC basic mechanisms modifying the statistical characteristics of a Data flow, in order to obtain a Random Bit Rate, CBR (constant bit rate) traffic.



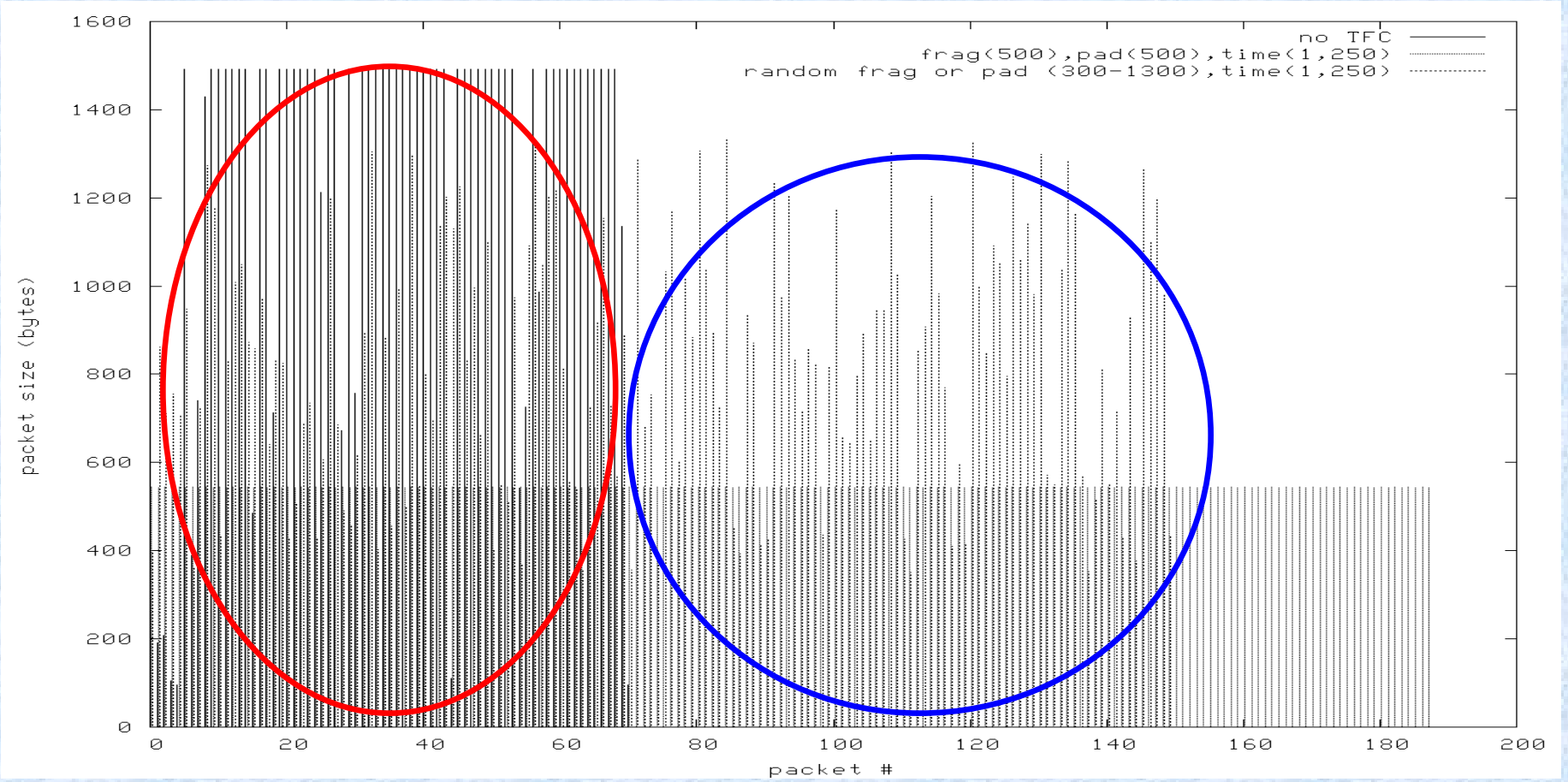


Protocol fingerprinting

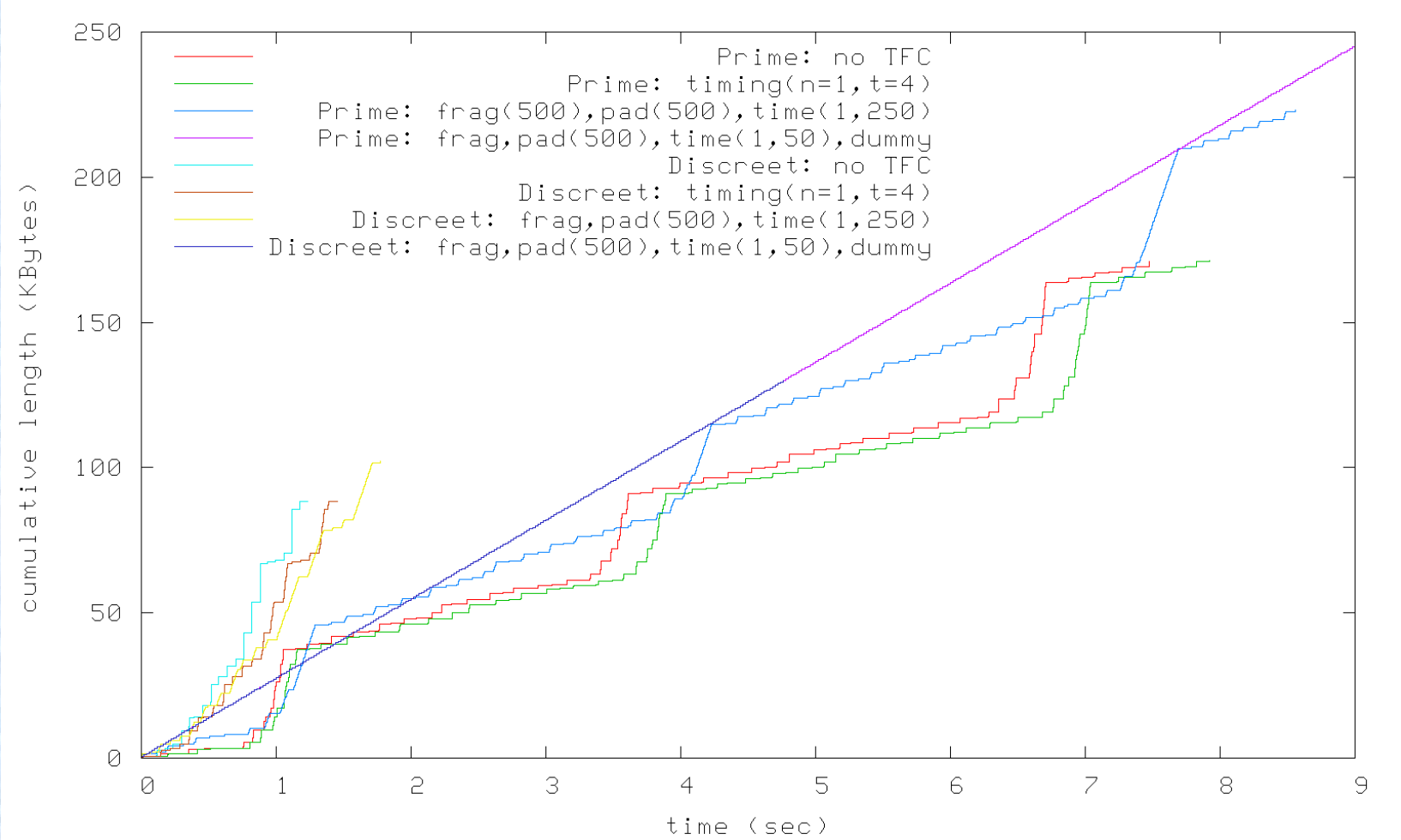
- Accurate flow classification exploit its very first packets
  - Length (L. Bernaille, R. Teixeira, and K. Salamatian, “Early Application Identification”, Proceedings of The 2nd ADETTI/ISCTE CoNEXT Conference, Portugal, 2006)
  - Inter-arrival time (M. Crotti, F. Gringoli, P. Pelosato, L. Salgarelli, “A statistical approach to IP-level classification of network traffic”, IEEE ICC 2006, 11-15 Jun. 2006)
- TFC tunnels avoid classification since
  - Packets are padded
  - Delay algorithms modify packets inter-arrival time
  - Different application flows can be multiplied on the same TFC SA.

Flows correlation

- The Discreet page downloads in 1.3 seconds and generates 88 Kbytes of traffic. The same download with CBR TFC takes 4.7 seconds and 130 KBytes



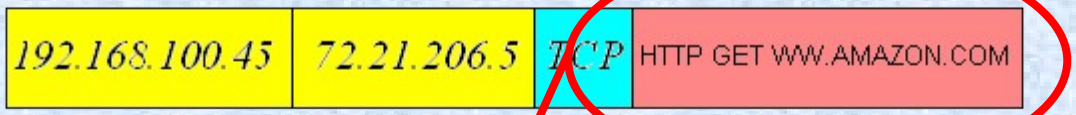
Web site fingerprinting



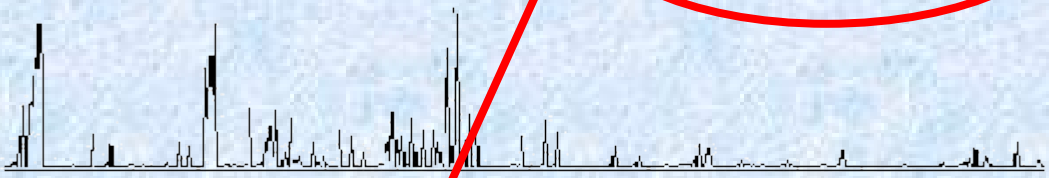
Conclusion

- Il codice è disponibile su
  - http://minerva.netgroup.uniroma2.it/discreet
- Per domande,suggerimenti etc
  - simone.teofili@uniroma2.it

Malicious Traffic Analysis



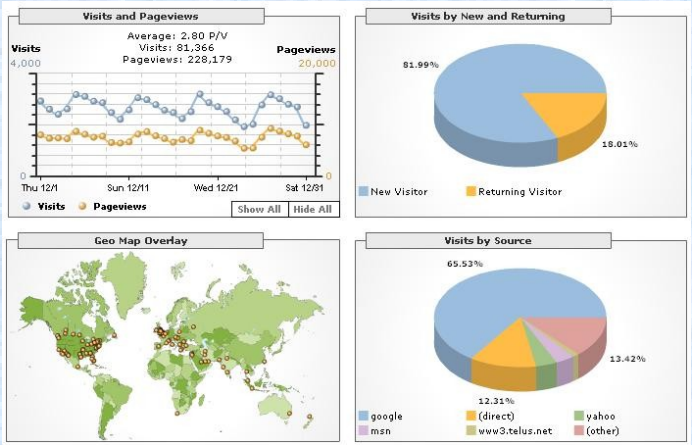
client
192.168.100.45



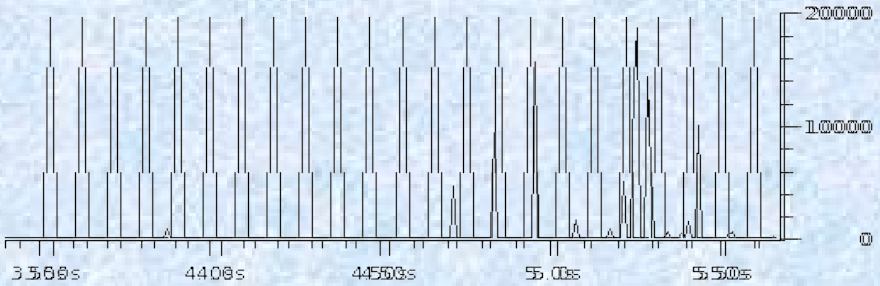
www.amazon.com
72.21.206.5

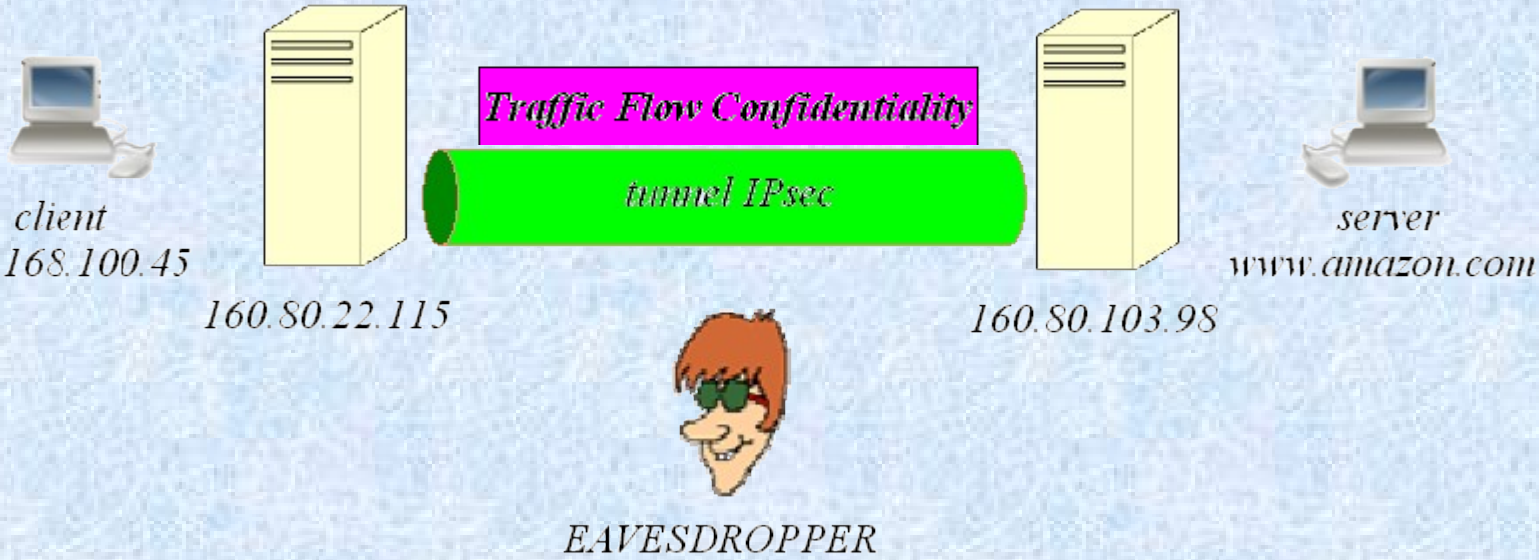
EAVESDROPPER

Download from Amazon

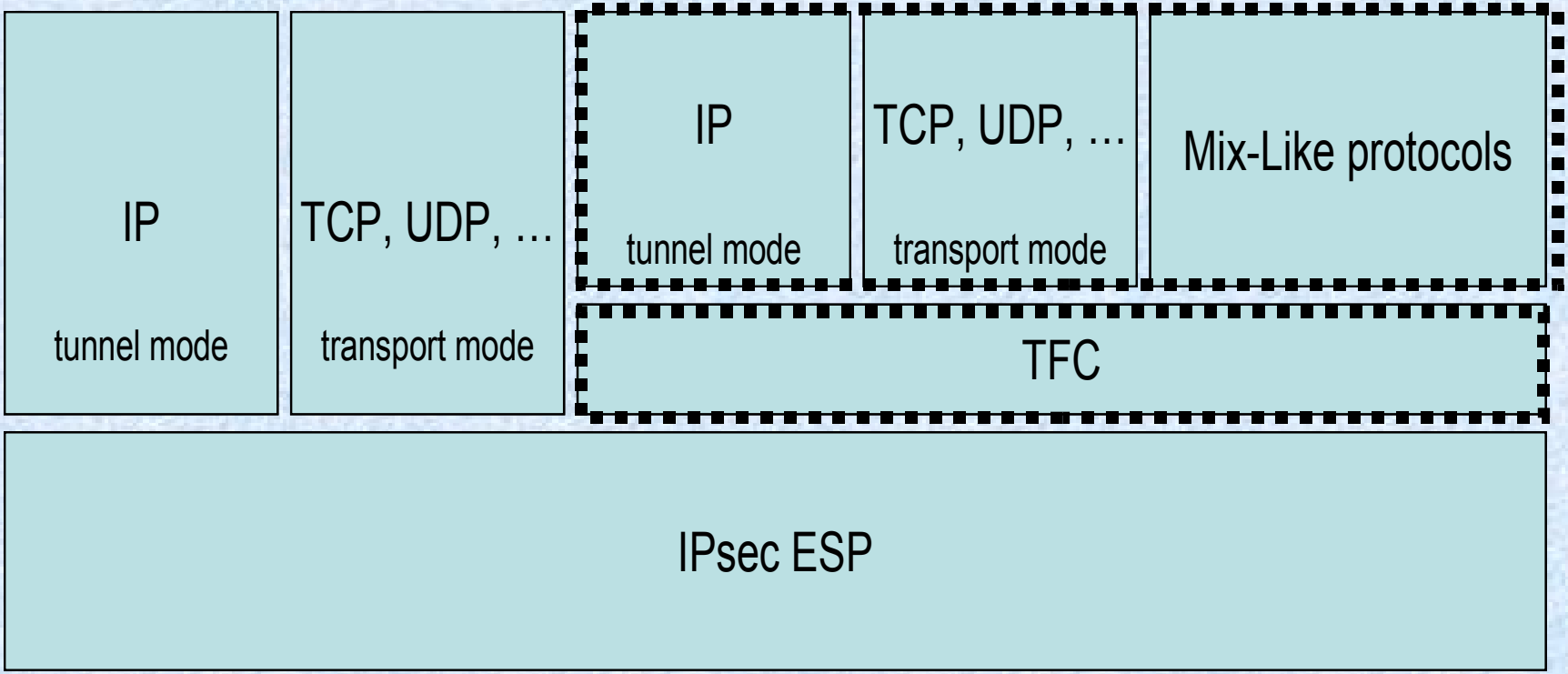


Traffic Flow Confidentiality

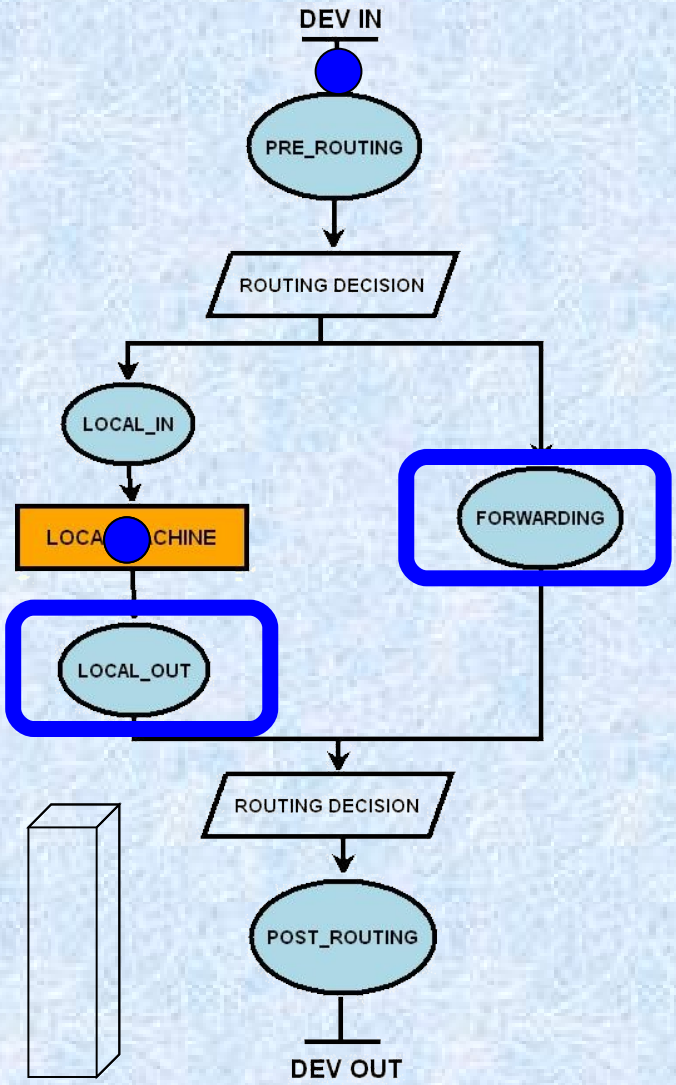




Traffic Flow Confidentiality



Output Stack



Dummy packets

- A timer is associated to each queue. When the timer expires, a packet from the head of the queue is sent and the next timer is set
- If the queue is empty, we create a new dummy packet (IP protocol = 59) and send it
- Since the queue is situated before IPsec encryption, dummy packets are sequentially encrypted with data packets