The re-ECN Protocol

ITG Fachgruppentreffen 5.2.3

Mirja Kühlewind
mirja.kuehlewind@ikr.uni-stuttgart.de
15.04.2010

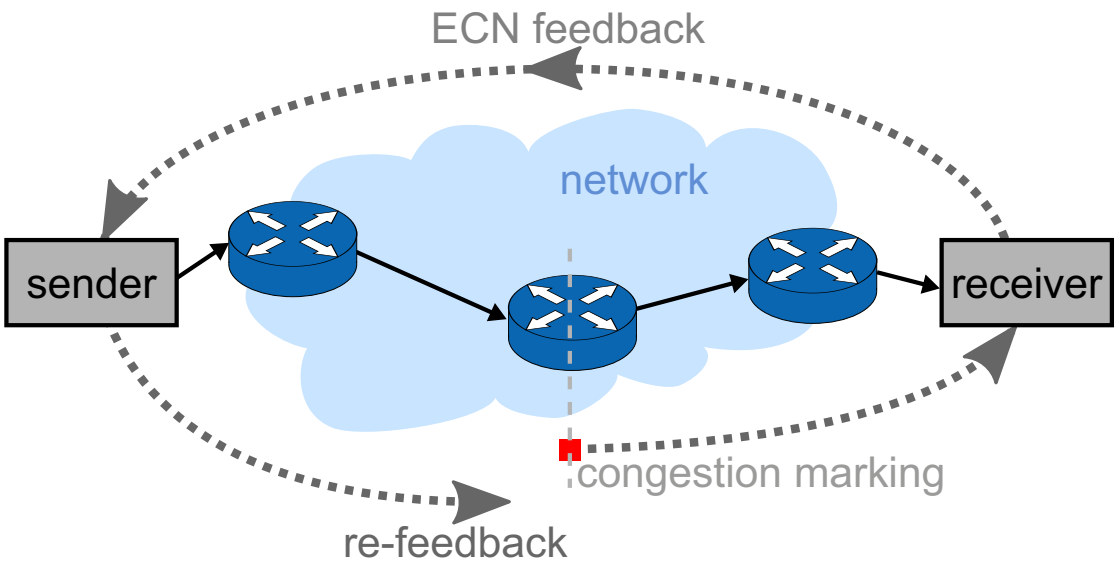
Universität Stuttgart
Institut für Kommunikationsnetze
und Rechnersysteme (IKR)
Prof. Dr.-Ing. Andreas Kirstädter

Overview of the re-ECN Protocol

A TCP/IP signaling mechanism for "Congestion Exposure"

Principle Protocol Mechanism

- Receive congestion information from receiver (Explicit Congestion Notification – ECN)
- Re-insert congestion information (= estimation about expected congestion level)



Goal of Congestion Exposure

- Monitor upstream/downstream congestion in the network and at network borders
- Make end-systems accountable for the congestion they cause in the network

Outline

- Problem Statement
- The re-ECN Protocol
- The re-ECN Framework
- Simulations with Linux Kernel Code
- Conclusion and Outlook

Problem Statement

Capacity Sharing in the Internet

The end-to-end principle

- Congestion control mechanisms in the end-systems avoid a congestion collapse
- Over-provisioning of network resources encounters workload peaks

Problems

- Amount of traffic in the Internet is increasing vs. decreasing profit/bit
  - More and more data though P2P, video streaming, network data storage
  - New services with challenging demands (const. data rate, interactive, high data volumes)
  - Higher access rates, always-on contracts
- Upgrading of the physical network will not improve Quality of Experience (QoE)
  - Aminority of heavy users allocate a large share of the bandwidth capacity
  - This traffic bothers time-critical transmission of other users
  - Physical expansion implys costs for all costumers, but only a minority benefits

→ ISP's do not want to extent their networks but want to achieve a fairer share

Problem Statement

Congestion Management in the Network

Current Practise

1. Application rate limiting for high bandwidth services (Deep Packet Inspection - DPI)
2. Volume counting at network ingress to improve the Quality of Experience (QoE) for the majority of the users
3. "Comcast's Protocol Agnostic Congestion Management System"

   → assign lower priority status to heavy users

Problems

1. IPSs do not know the end-systems intent

   → application rate limiting is not appropriate (Network neutrality, Fairness)
2. ISPs do not know the congestion situation in their network

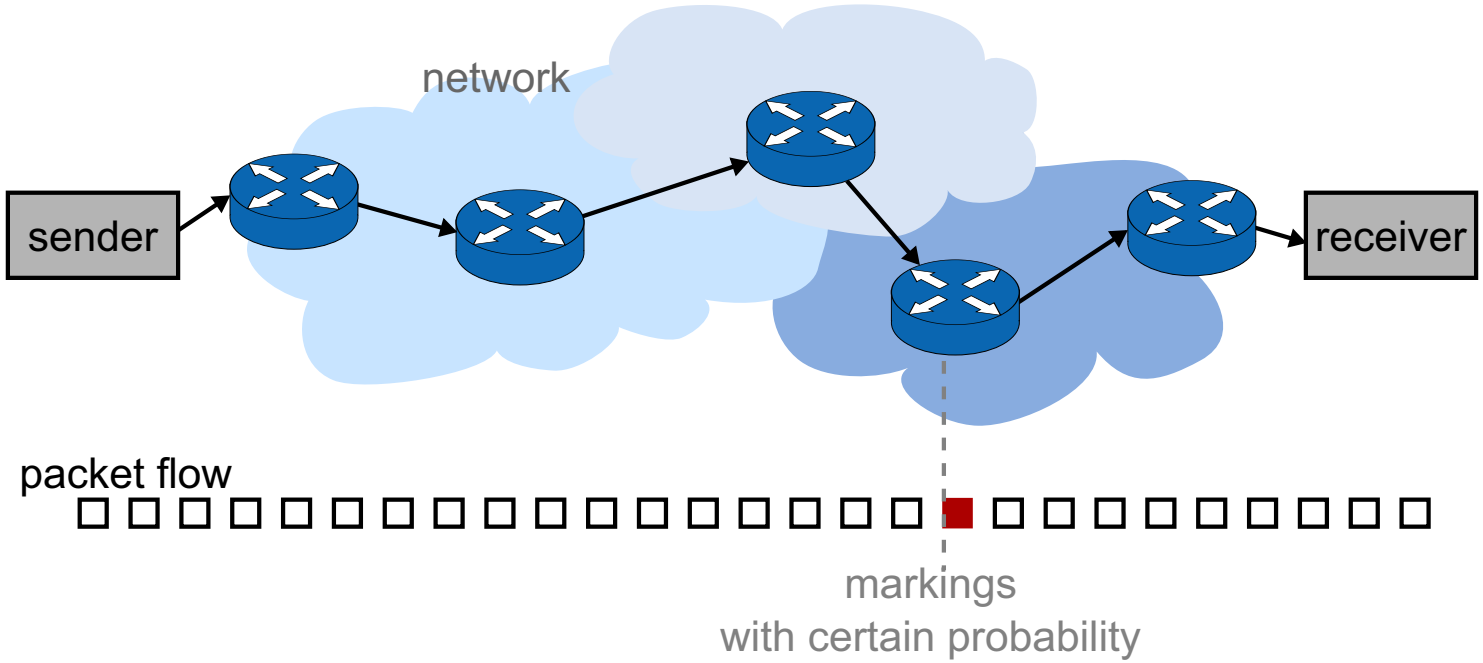
   → policing would only be necessary if congestion occurs

   → policing would be most beneficial at network ingress

→ End-system-based congestion control with per-user fairness (Congestion Accounting)

The ECN Protocol

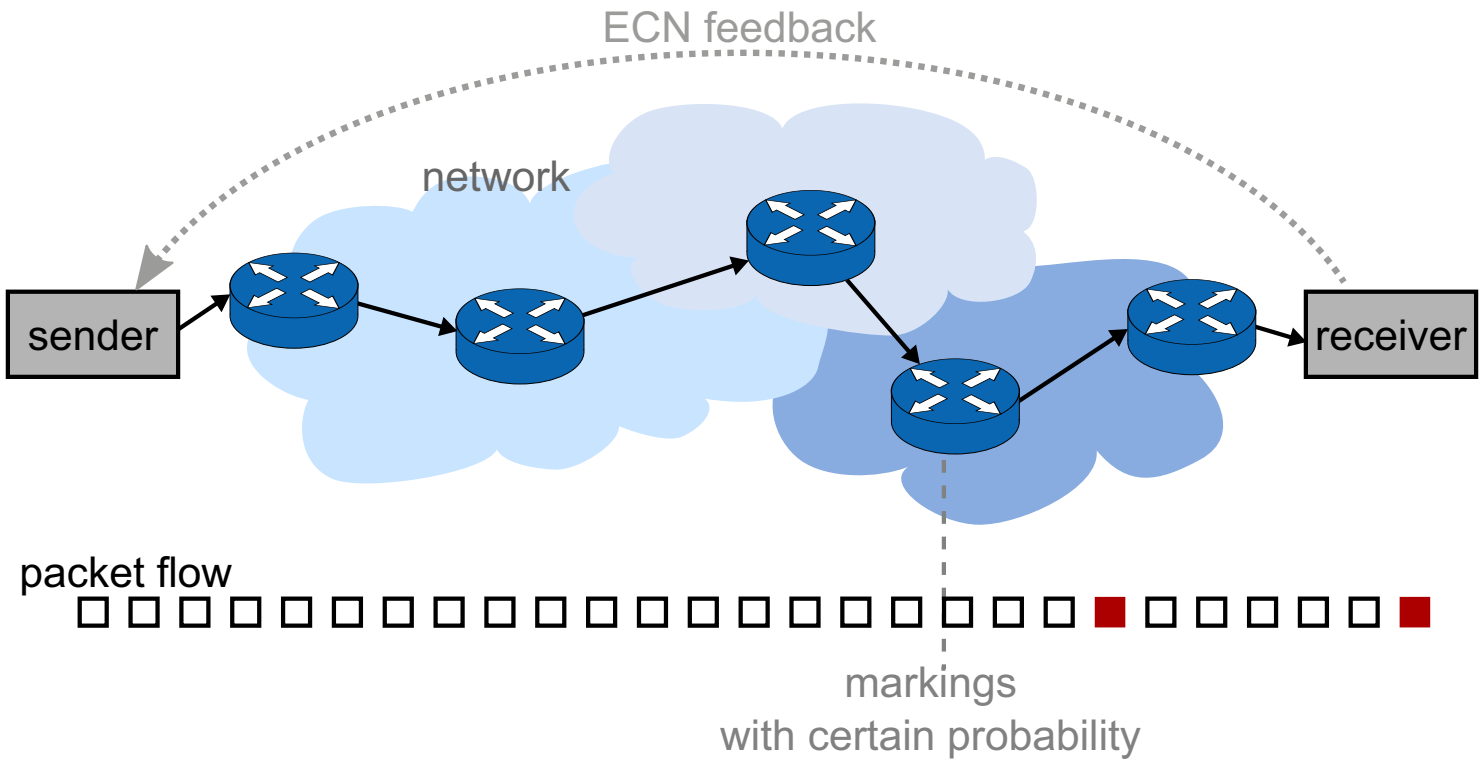
ECN (Explicit Congestion Notification)



1. Routers mark packets (instead of dropping them – Random Early Detection, IP flag)

The ECN Protocol

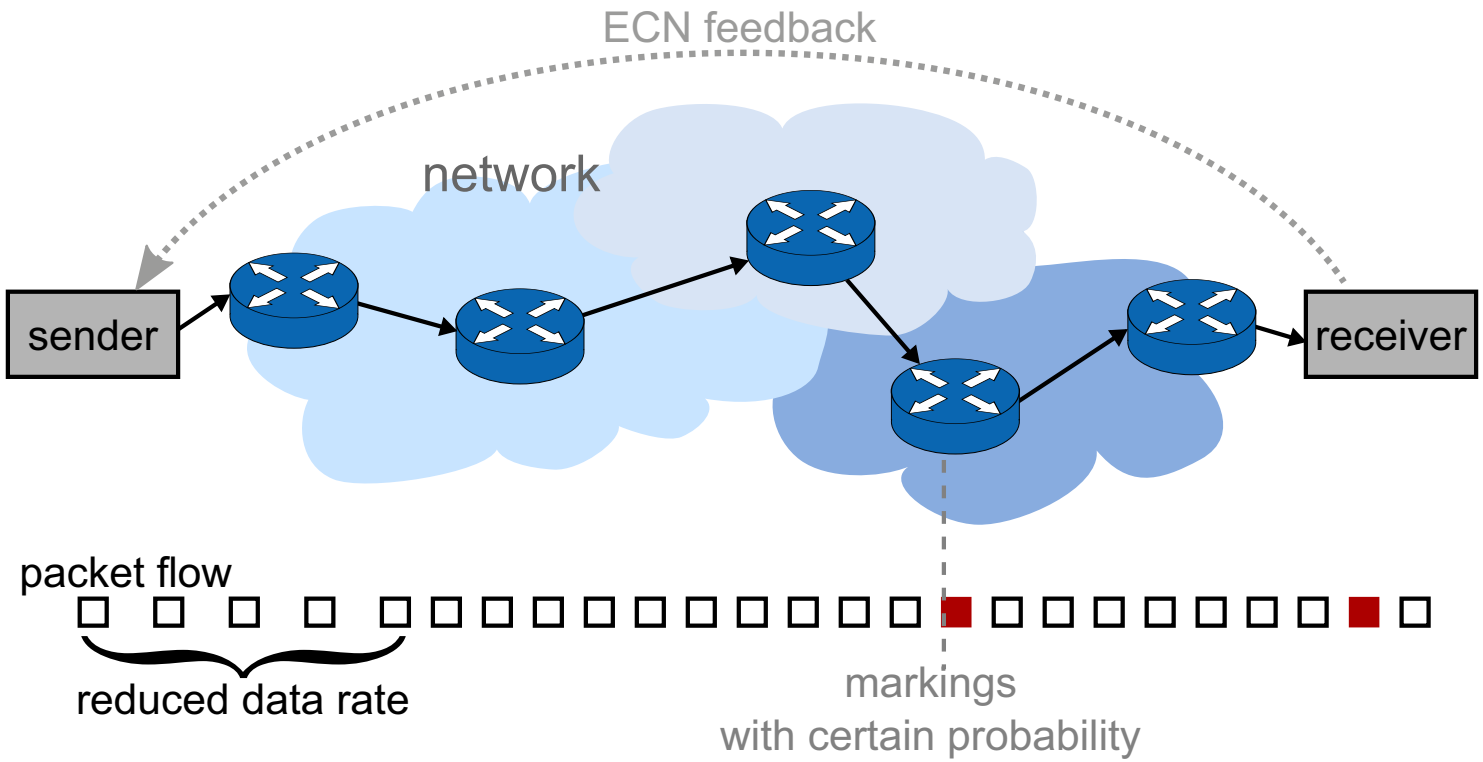
ECN (Explicit Congestion Notification)



1. Routers mark packets (instead of dropping them – Random Early Detection, IP flag)
2. ECN receiver feeds congestion announcement back to the sender (TCP ACK)

The ECN Protocol

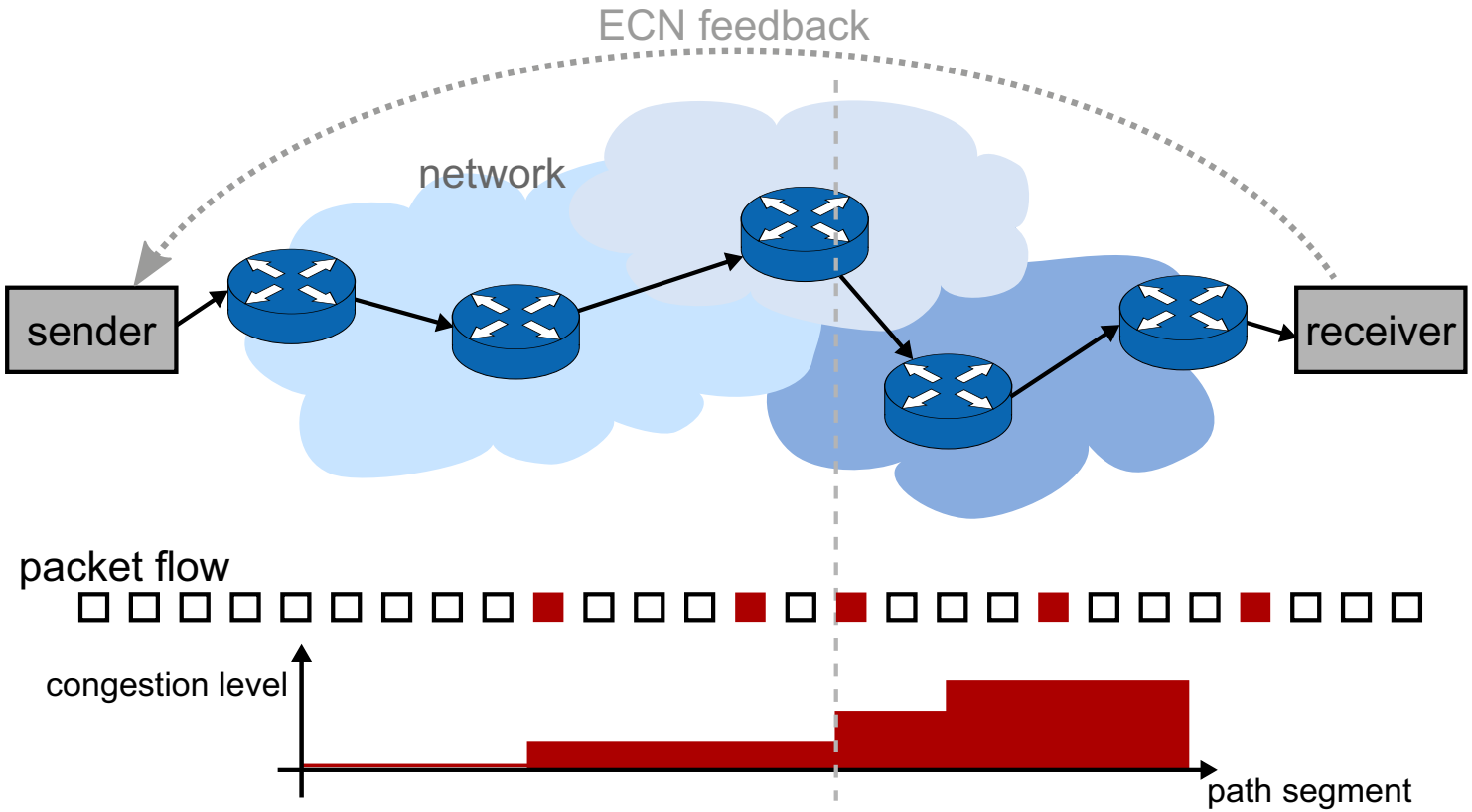
ECN (Explicit Congestion Notification)



1. Routers mark packets (instead of dropping them – Random Early Detection, IP flag)
2. ECN receiver feeds congestion announcement back to the sender (TCP ACK)
3. ECN sender reduces congestion windows (as on loss – TCP Congestion Control)

The re-ECN Protocol

re-ECN (re-insert ECN) – expose expected whole path congestion to network elements

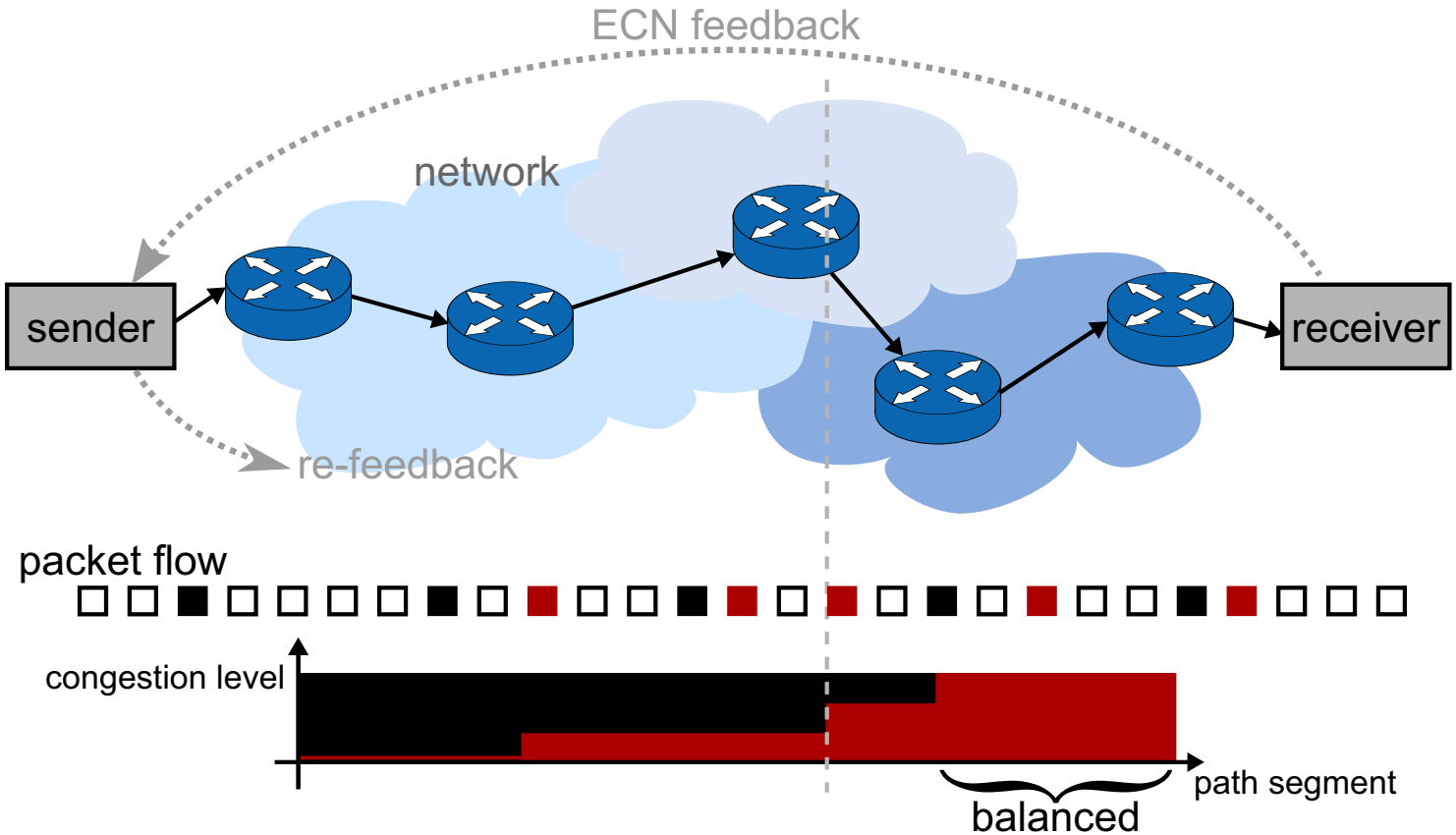


re-ECN sender marks a packet for every congestion announcement from the receiver

- Fraction of red marks ("congested") gives level of congestion experienced so far

The re-ECN Protocol

re-ECN (re-insert ECN) – expose expected whole path congestion to network elements

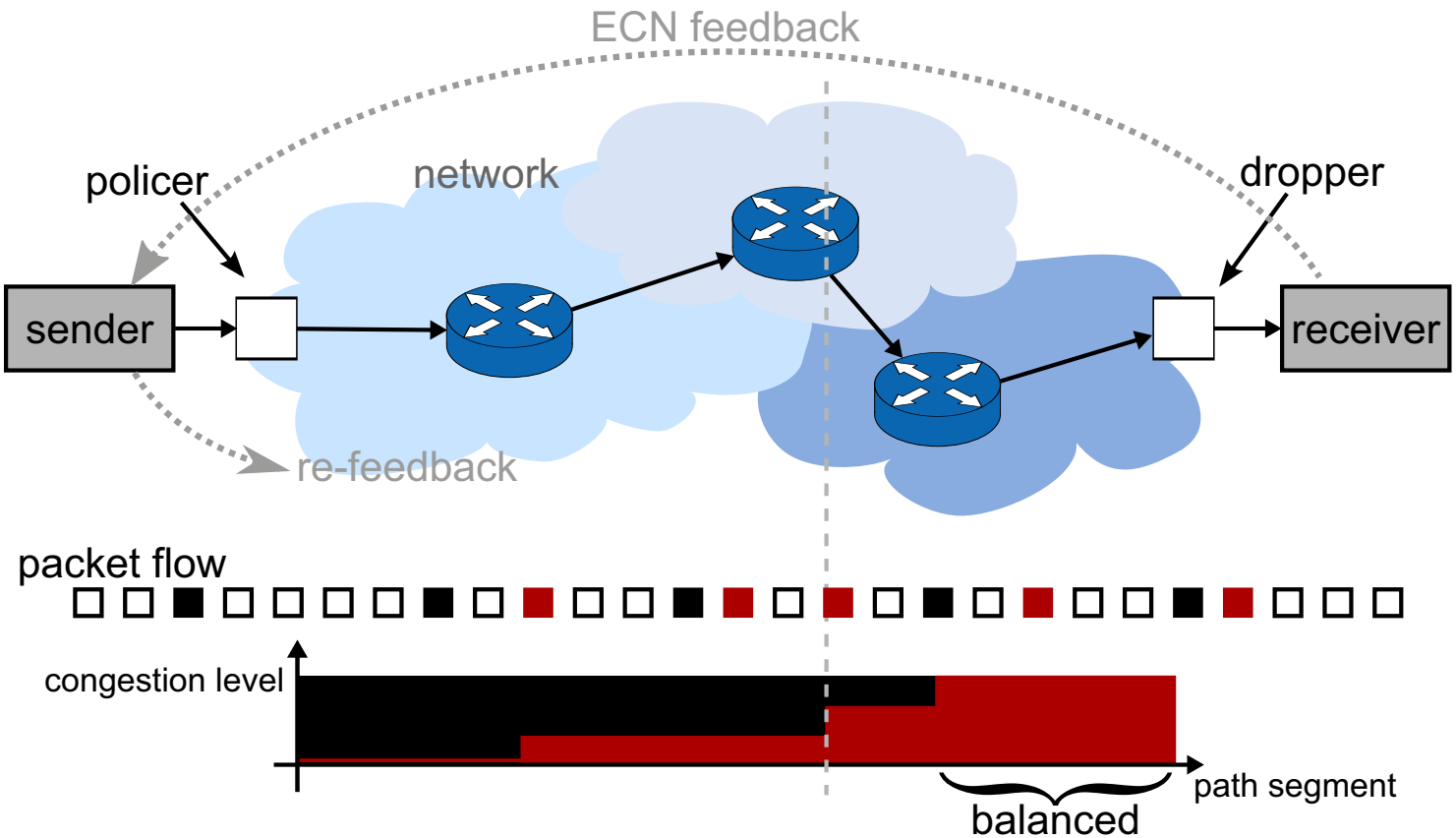


re-ECN sender marks a packet for every congestion announcement from the receiver

- Fraction of red marks ("congested") gives level of congestion experienced so far
- Fraction of black marks ("congestion expected") give the whole path congestion

The re-ECN Framework

Congestion Accounting – enable appropriate congestion control



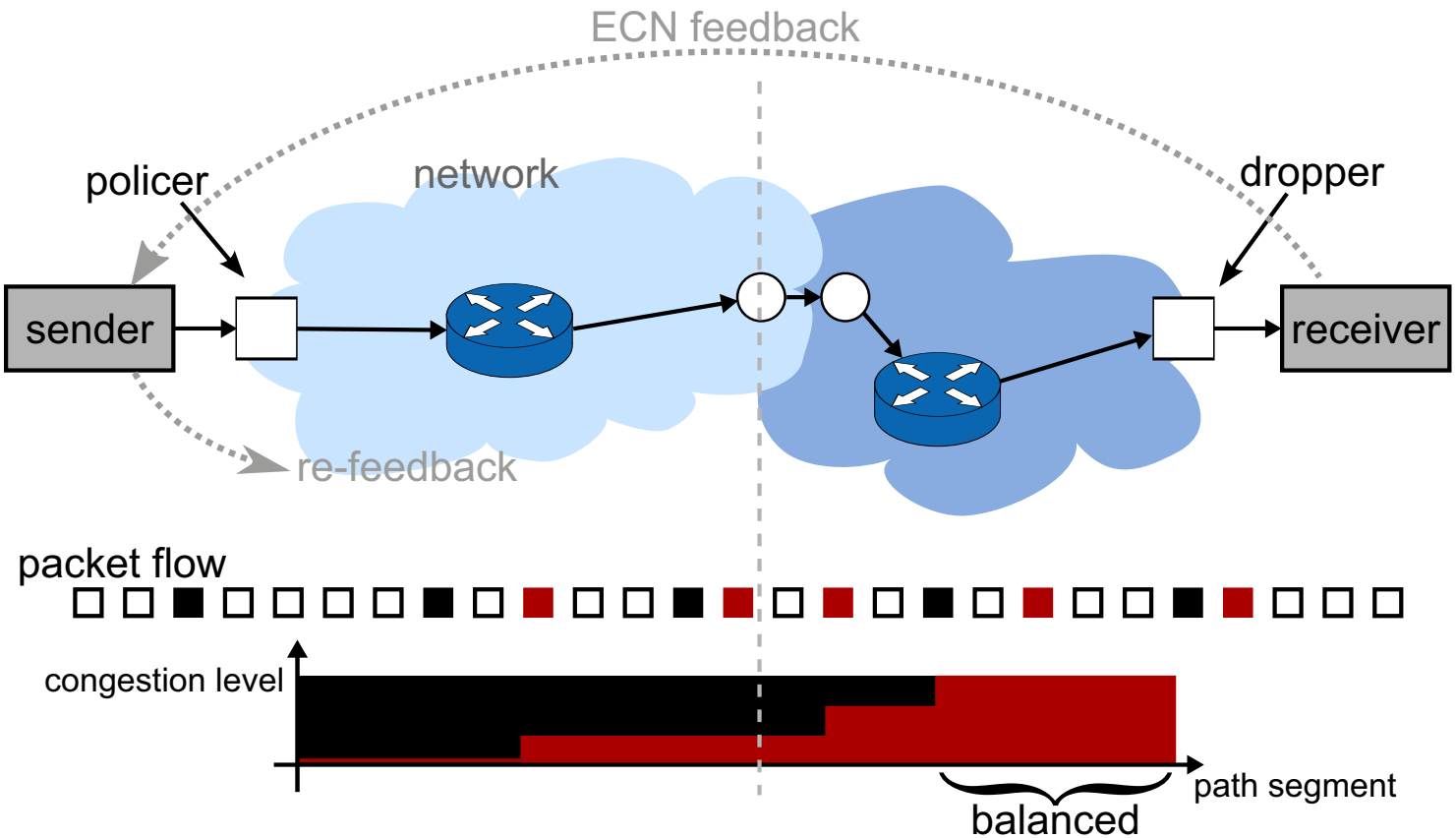
Dropper Detection of permanent negative flows

Policer Limitation of congestion volume (pos. marked packets, Token-Bucket principle)

Inter-Domain Policer (Border-charging, flow control, routing)

The re-ECN Framework

Inter-Domain Policer – border charging and inter-domain (re-)routing



Dropper Detection of permanent negative flows

Policer Limitation of congestion volume (pos. marked packets, Token-Bucket principle)

Inter-Domain Policer (Border-charging, flow control, routing)

Congestion Accounting with re-ECN

Exposure of downstream congestion to network components

→ can be used for inter-domain SLAs

→ may be used for traffic management in the network

Congestion Exposure at network ingress

→ establishes an information equilibrium between end-system and network

→ enables to police packets before imposing congestion

→ permits to use as much as possible of the available capacity (if no congestion occurs)

Per-costumer limitation of congestion volume

→ introduces an incentive to not cause more congestion than needed while sending data

→ provides the basis for a fair share between costumers (not flows or applications or...)

Network Simulation with Linux kernel code

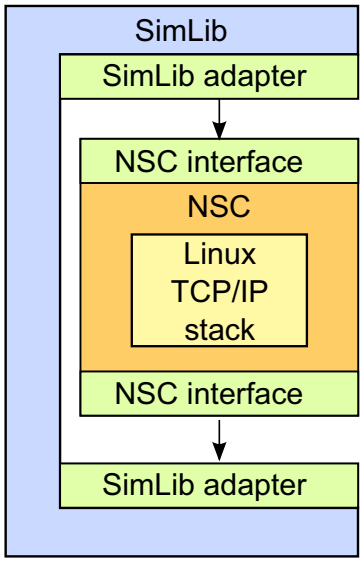
IKR SimLib and Network Simulation Cradle

Network Simulation Cradle (NSC)

http://research.wand.net.nz/software/nsc.php

"cradles real world operating systems' network stacks in a wrapper"

- Framework which allows Linux, FreeBSD, OpenBSD and lwIP code to be excecuted as user space program
- Provides interface to be used in a network simulation context



IKR SimLib

http://www.ikr.uni-stuttgart.de/en/Content/IKRSimLib/

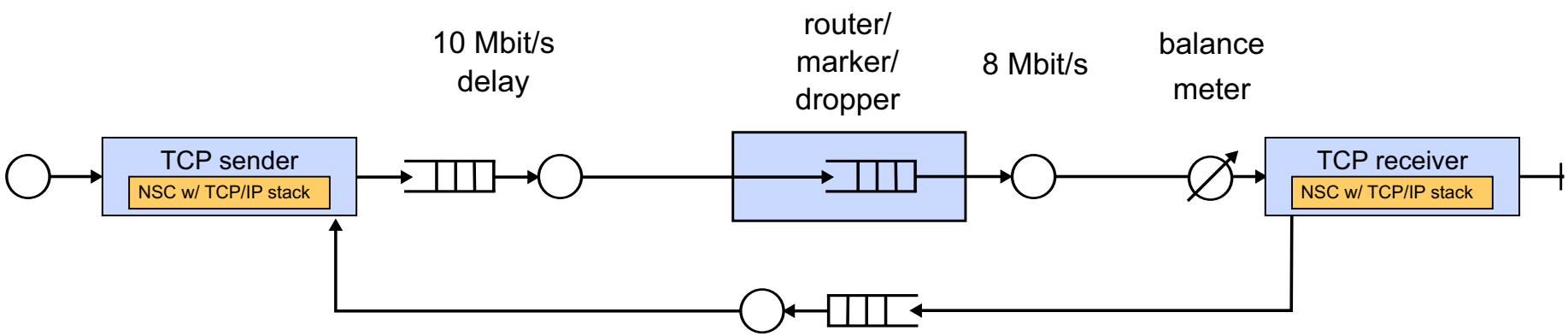
- Event-based **Sim**ualtion **Lib**rary of the Institute of Communication Networks and Computer Engineering, University Stuttgart
- NSC TCP adapter allows to measure realistic traffic traces in complex scenarios
  - Copy function allows easy configuration of complex topologies
  - Generation of syntetic traffic or replay of Internet traffic traces

Network Simulation with Linux kernel code

Simulation Scenario with Linux kernel code

Functional analysis of the re-ECN protocol

- Implementation: TCP/IP-Stack Linux Kernel 2.6.26
- Simulation: SimLib with Network Simulation Cradle (version 0.5.0)

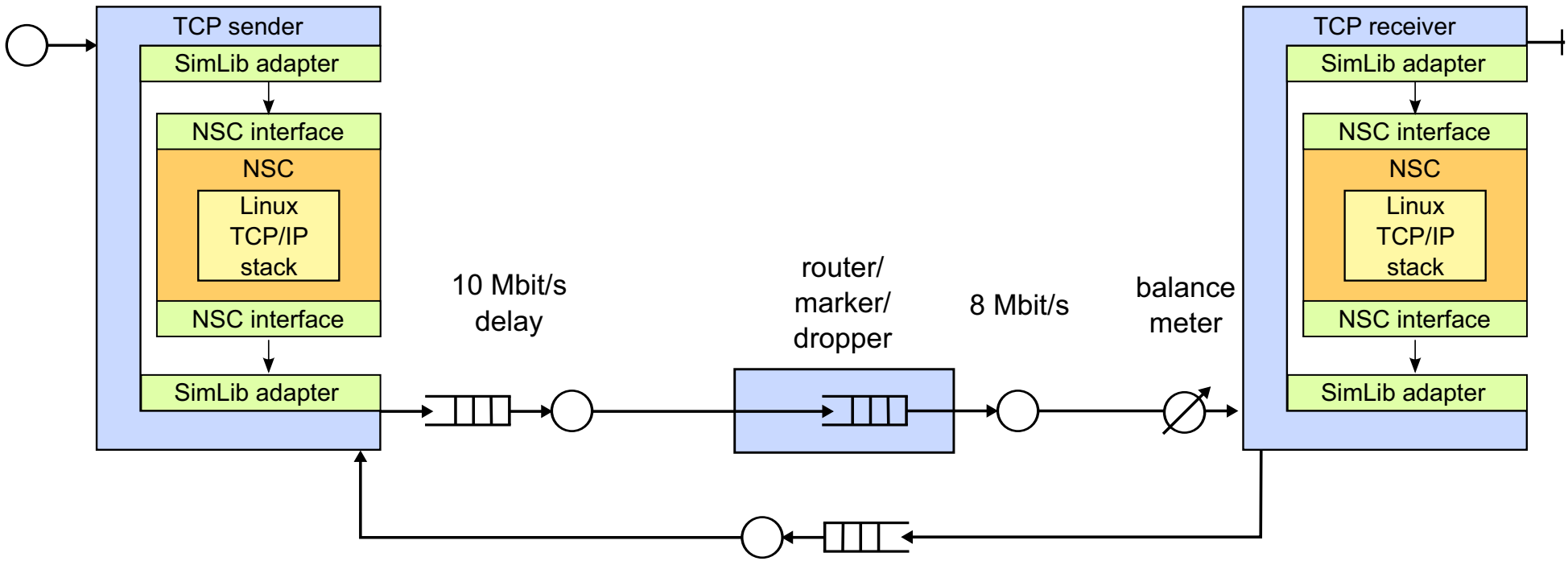


Network Simulation with Linux kernel code

Simulation Scenario with Linux kernel code

Functional analysis of the re-ECN protocol

- Implementation: TCP/IP-Stack Linux Kernel 2.6.26
- Simulation: SimLib with Network Simulation Cradle (version 0.5.0)

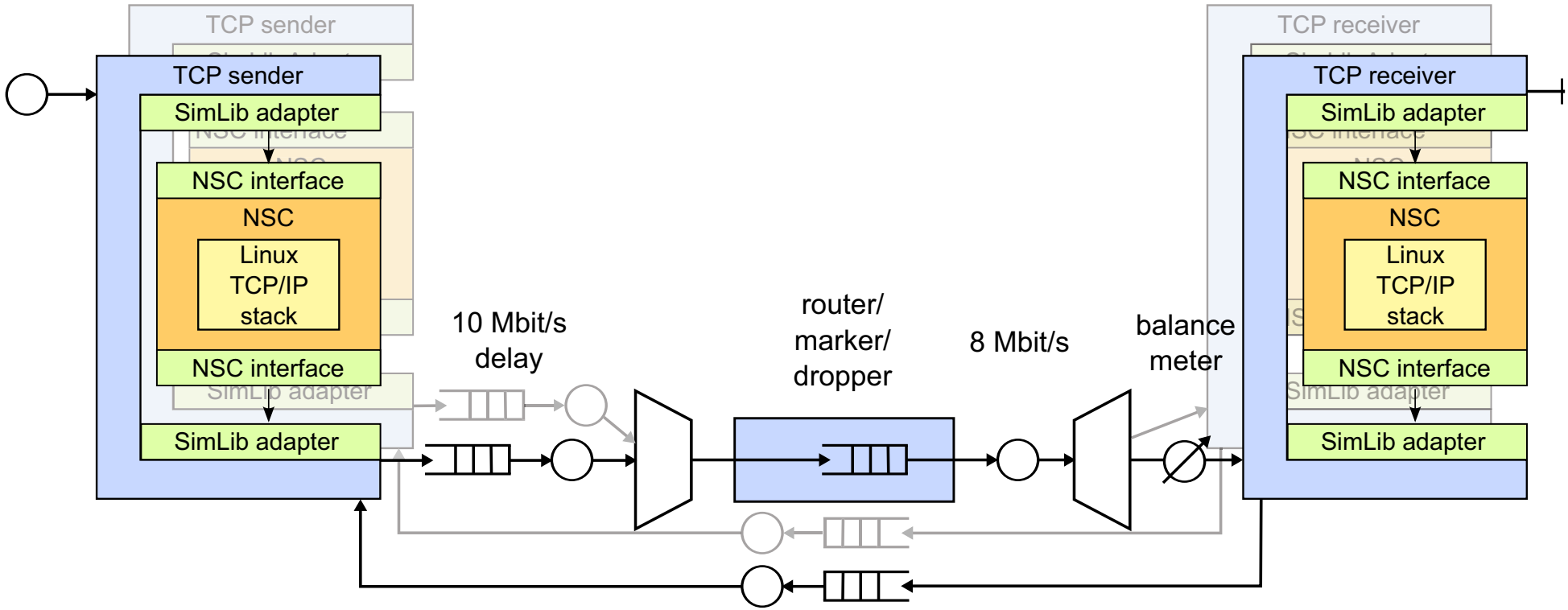


Network Simulation with Linux kernel code

Simulation Scenario with Linux kernel code

Functional analysis of the re-ECN protocol

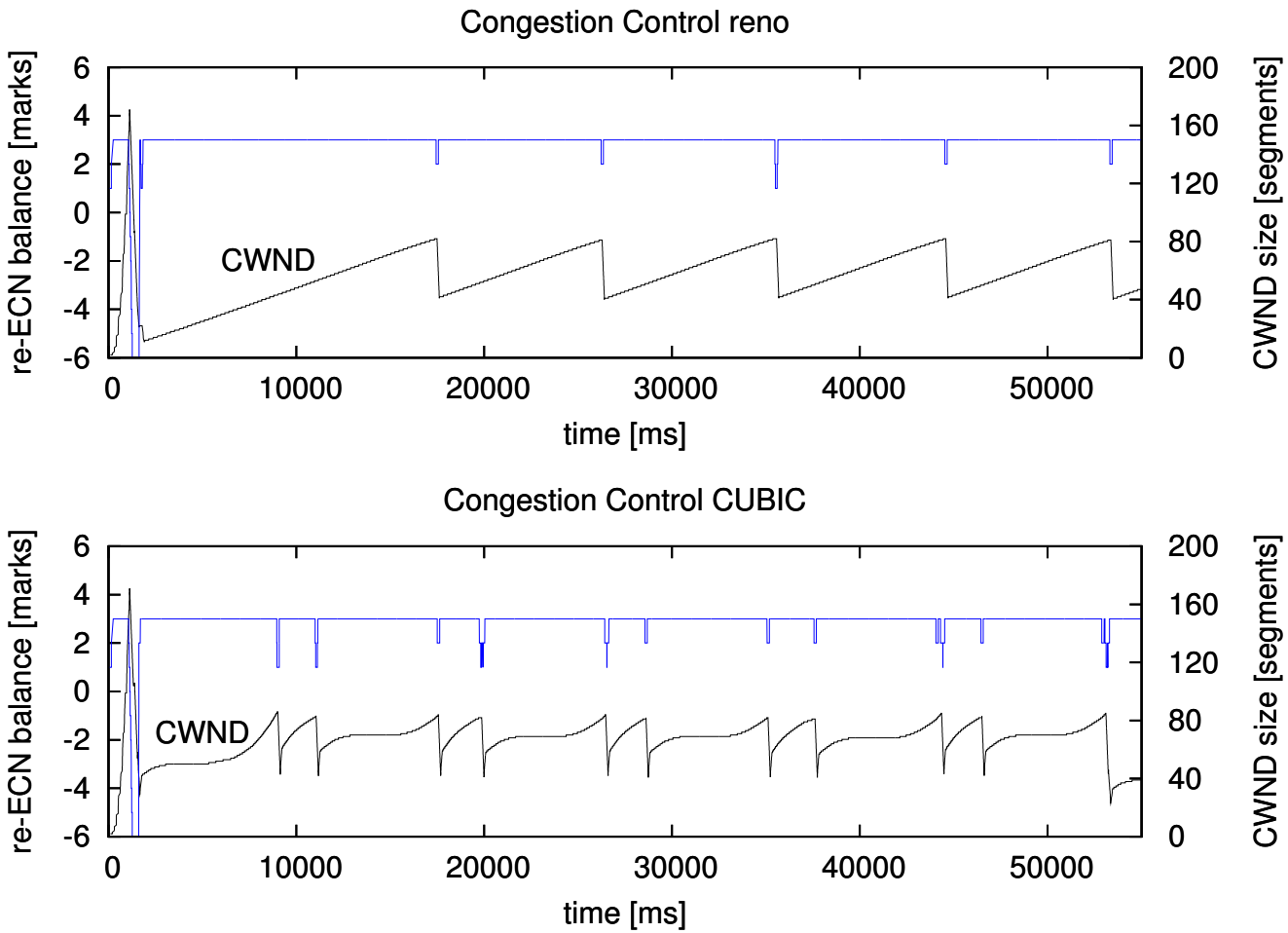
- Implementation: TCP/IP-Stack Linux Kernel 2.6.26
- Simulation: SimLib with Network Simulation Cradle (version 0.5.0)



Network Simulation with Linux kernel code

Congestion Exposure with the re-ECN Protocol

re-ECN balance of one TCP connection at network egress with bulk traffic



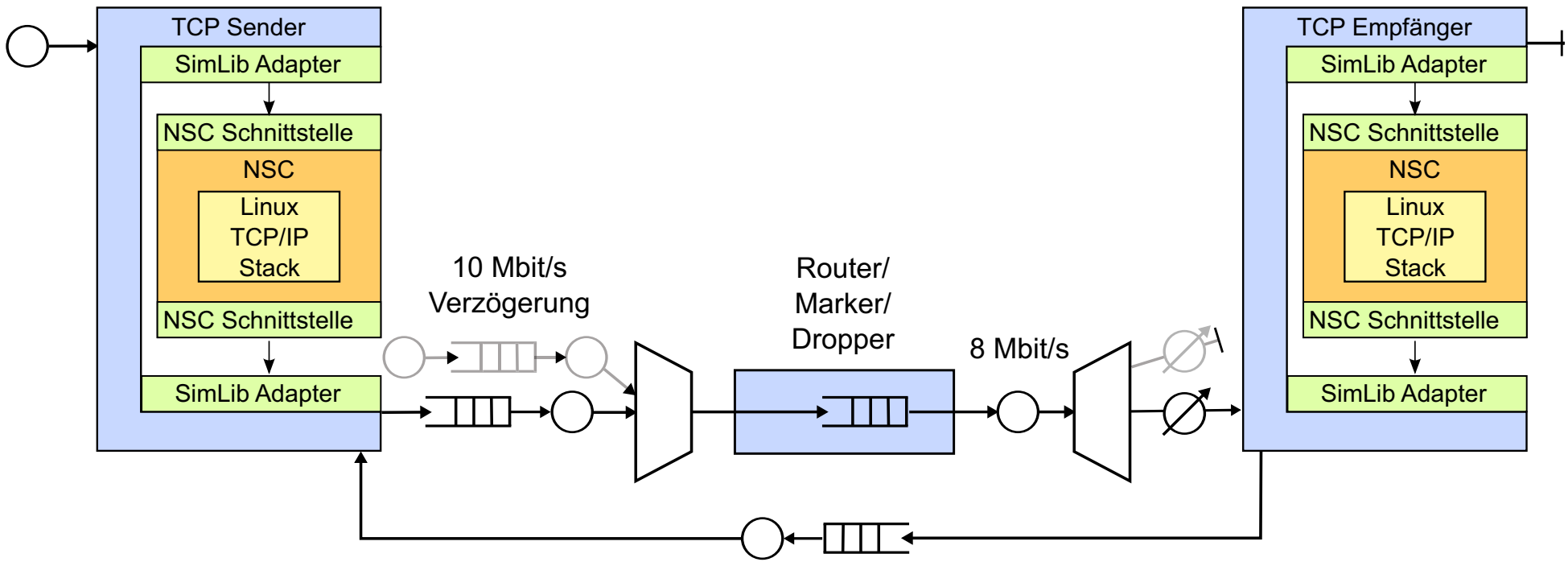
→ Number of congestion marking differs but is relevant for Congestion Accounting

Network Simulation with Linux kernel code

Simulation runs with const. bit rate (CBR) cross traffic

Scenario

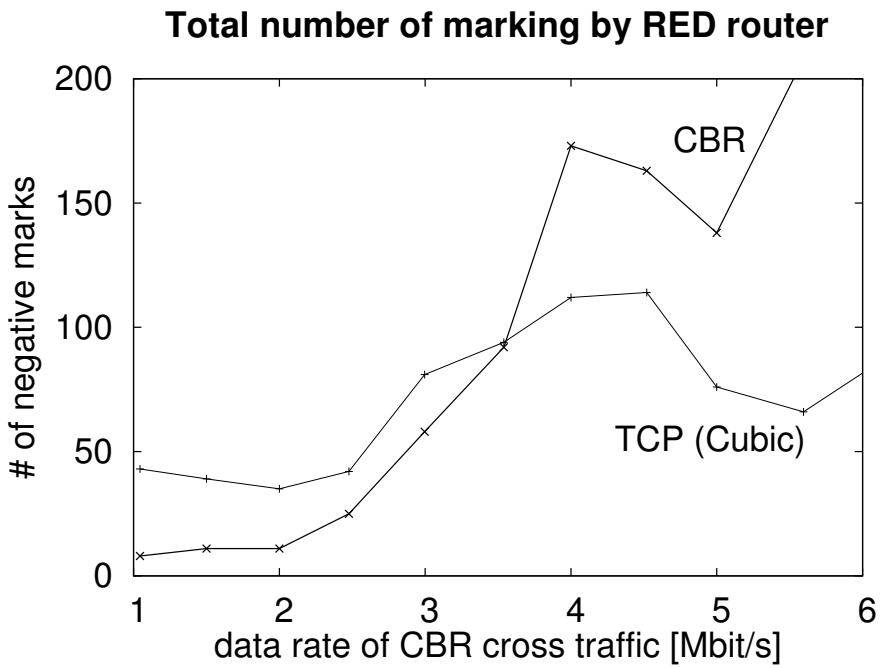
- Cross traffic with constant bit rate and greedy source
- Assume positive marked packet 100ms (min. RTT) after a negative mark for CBR traffic
- Several simulation runs with cross traffic with different CBRs (1 minute each)

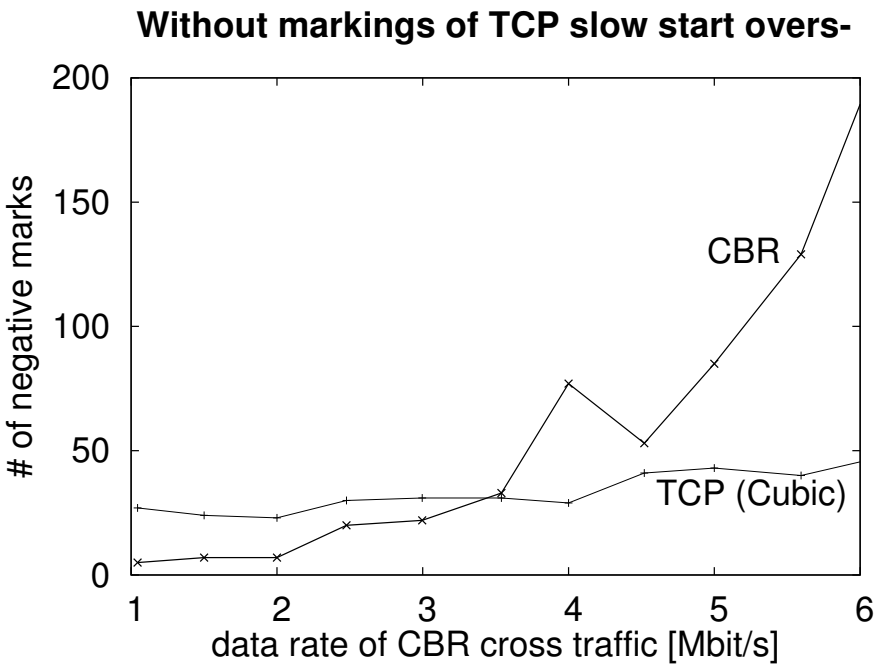


Network Simulation with Linux kernel code

Preliminary Results

Number of negative markings with const. bit rate (CBR) cross traffic (8Mbit/s capacity)





- Number of markings depends on share of channel capacity and TCP CA periods
- Less markings for TCP flows in slow start overshoot with high CBR

Conclusion and Outlook

Summary

- re-ECN: Exposure of expected congestion on a network path (to network elements)
- Policer: Limitation of congestion volume per end-system/costumer

  → can enable new congestion control schemes
- Impact of path characteristics and other protocol mechanism on re-ECN
  - Different AQM/RED parameterization and congestion control/Slow-Start options influence the total number of markings for one flow under the same network conditions
  - Total number of markings is relevant for Congestion Accounting

    ↳ new protocol mechanisms are needed to minimize congestion markings

Open Issues

- Policer and dropper design
- Congestion management in the end-system
- Economic implications (Inter-domain charging and Inter-domain routing)