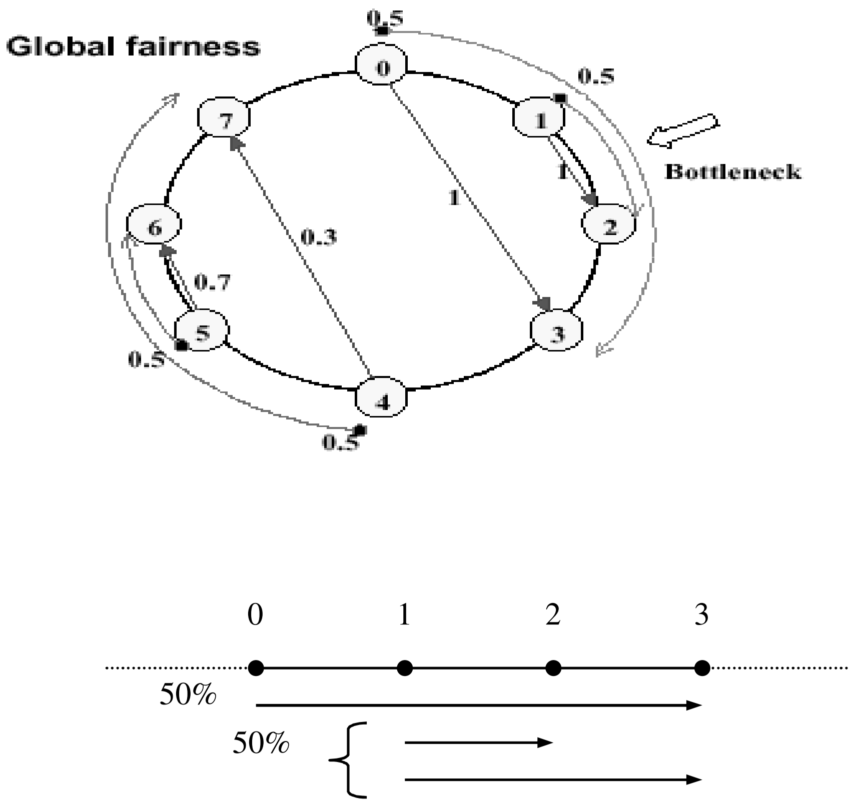
Metrics for Performance Evaluation

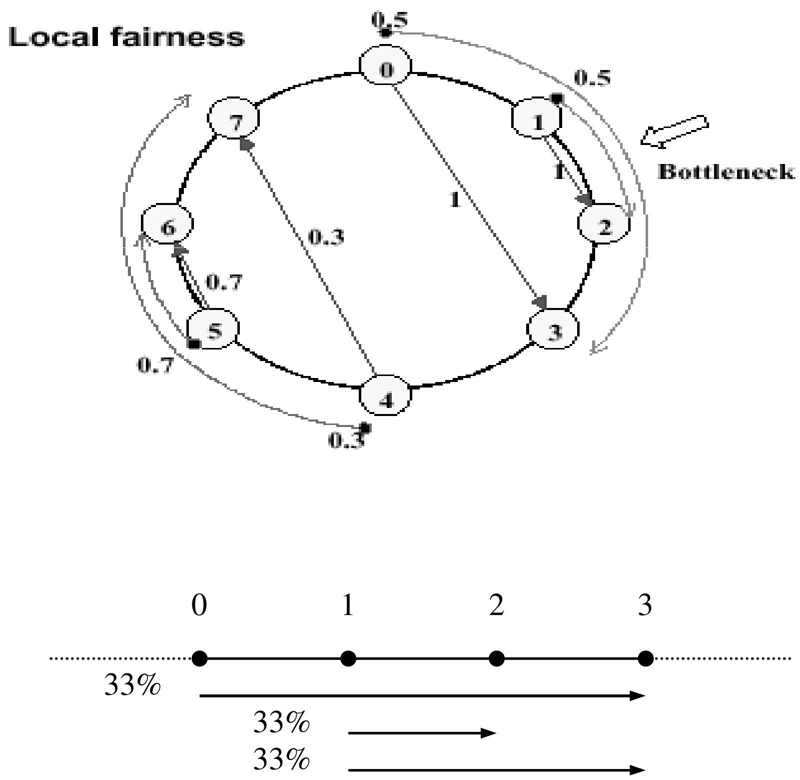
Nelson Fonseca
State University of Campinas

Metrics

- Express how well the systems is operating and achieving its goals, either the system as a whole or specific control mechanisms
- Should be measurable
- Needs to be relevant to the aimed goals
- Can be evaluated either as aggregate or per flow (connection, class) based

Metrics


Global fairness
0.5
0.5
Bottleneck
1
1
0.3
0.7
0.5
0.5
0
1
2
3
4
5
6
7
0
1
2
3
50%
50%


Local fairness
0.5
0.5
Bottleneck
1
1
0.3
0.7
0.7
0.3
0
1
2
3
4
5
6
7
0
1
2
3
33%
33%
33%

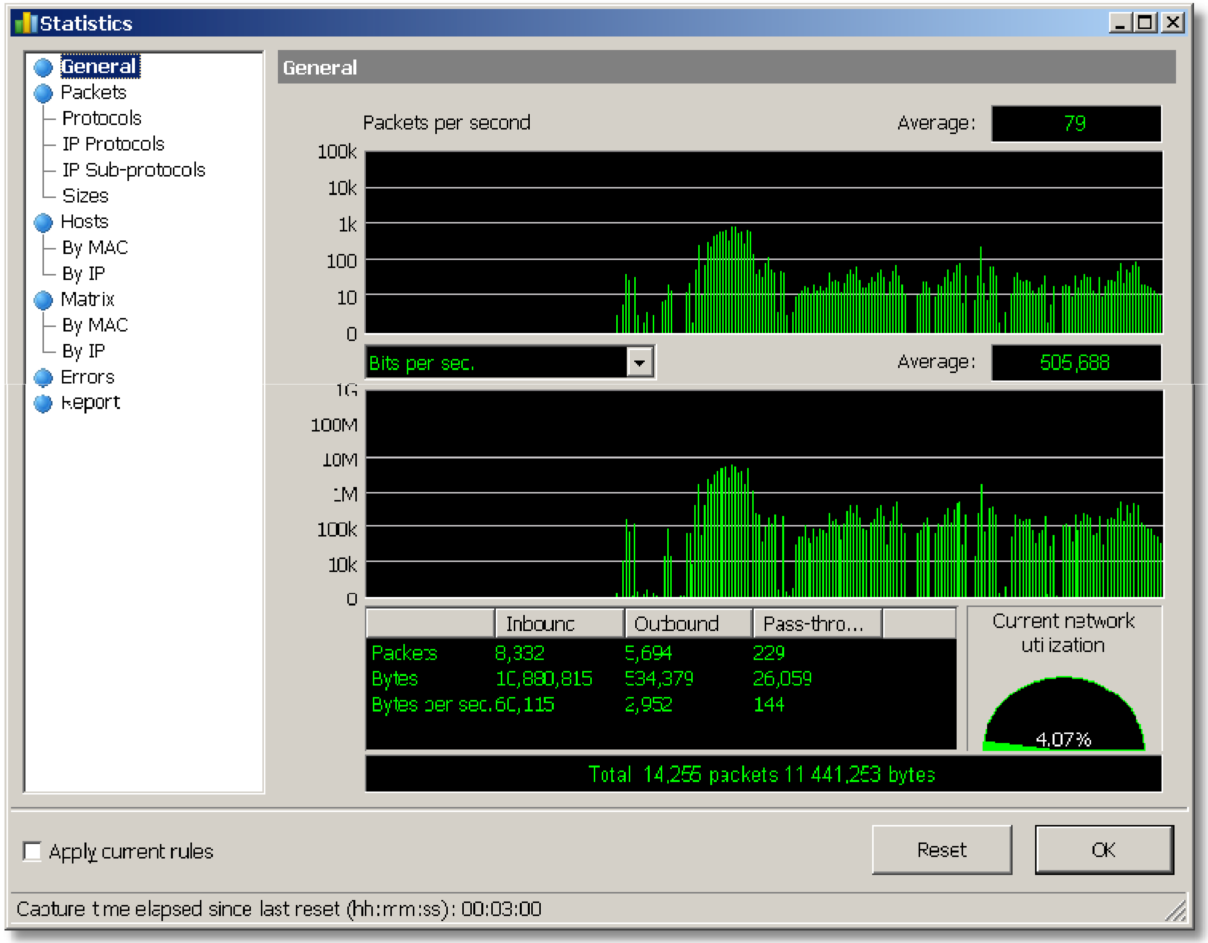
Metrics

- Most common:
  - Utilization;
  - Delay, jitter
  - Loss rate, loss burst
  - Blocking Probability
  - Troughput, Goodput
- Target to specific mechanisms
  - Response time
  - Convergence

Utilization

- Proportion of capacity use
- [0, 1]
- Offered load / capacity

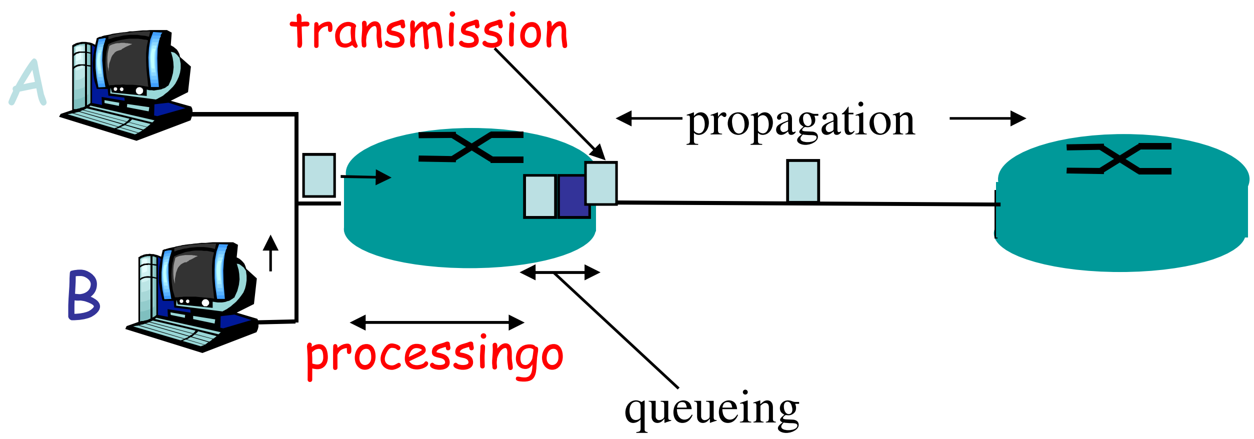
Utilization



Delay
Retardo, Atraso

- Elapsed time between the transmission of the first bit at the source and reception of the last bit of the packet at the destination
- Link delay – delay experienced by the packet to go through the link, including queuing delay to have access to the link
- End-to-end delay – expresses the delay between two end points that can be between transport level points, IP stations
- Round Trip Time (RTT)

Delay



Delay
Retardo, Atraso

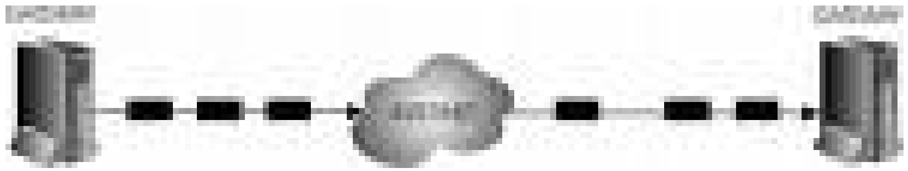
- Mean value usually does not express the distribution

- Needs to consider index of variation such as percentile

Jitter

- Variation between delays experienced
- Example the variation in the interarrival times of packet at the destination
- Impacts the quality of signal received, specially for real-time applications

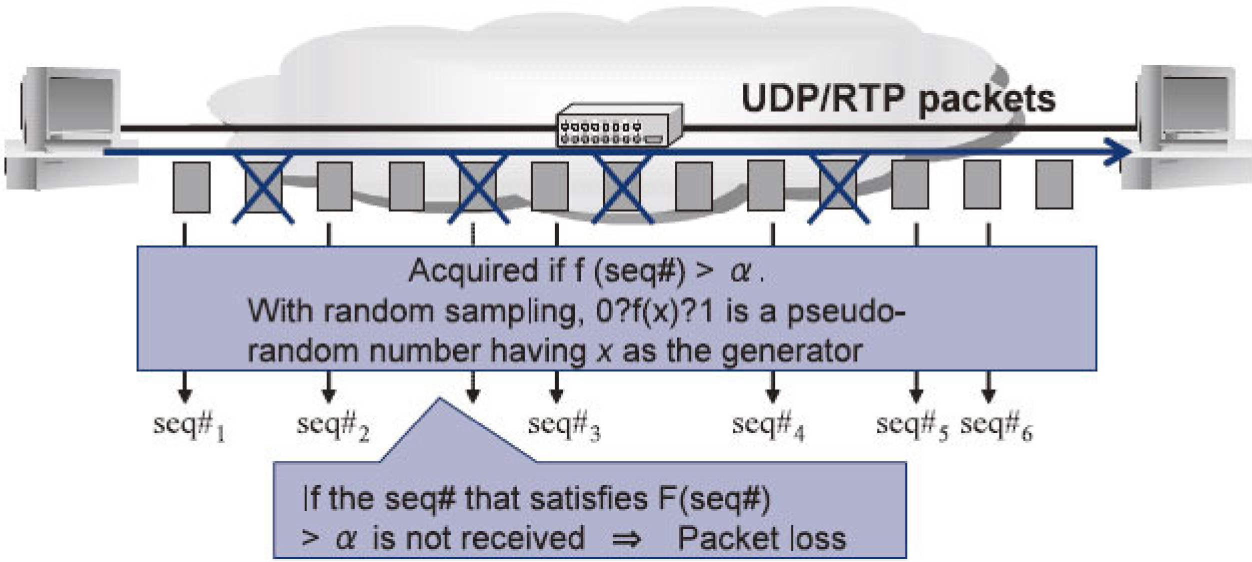
Jitter



Loss Rate
Taxa de Perda

- Expresses the proportion of packet lost
- Most precisely – loss ratio
- Loss probability – although used, not exactly correct due to existing correlations in traffic stream

Loss Rate



UDP/RTP packets
Acquired if f (seq#) > α .
With random sampling, 0?f(x)?1 is a pseudo-random number having x as the generator
seq#1
seq#2
seq#3
seq#4
seq#5
seq#6
If the seq# that satisfies F(seq#) > α is not received ⇒ Packet loss

Loss Burst, Gap loss

- Number of packet consecutively lost
- burst density - fraction of packet in bursts
- Duration of burst (mean duration of a burst loss), frequency of burst
- Impacts quality of recovered signal
- Example: 0.25 in 100 packets can be distributed one out of four or 25 losses in sequence

Blocking Probability
Probabilidade de Bloqueio

- Probability of not obtaining what is target at, usually access to a domain
- Examples: blocking probability in admission control, in circuit (path) establishment
- Can be given in Erlangs

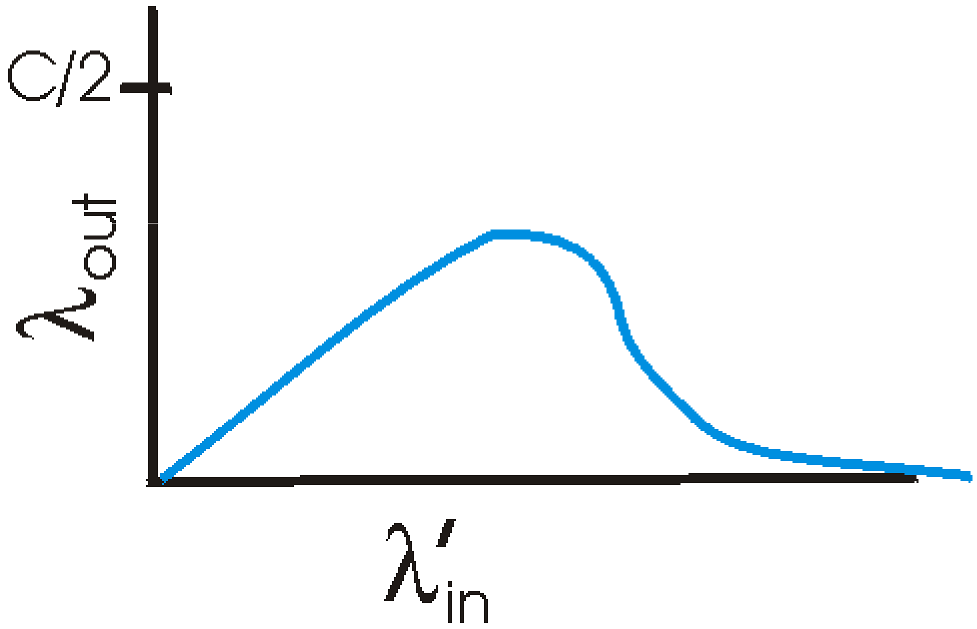
Throughput
Vazao

- Amount of bits transmitted per unit of time
- Includes all bits

Goodput

- Amount of useful bits transmitted per unit of time, does not include retransmission
- Expresses the usefulness of a scheme
- Throughput may be high while goodput low (high number of losses)

Goodput



Metrics for evaluation of Congestion Control Mechanisms

(currently under definition)

Response Times to Changes

- Response time to sudden network changes,
- Examples: changes in bandwidth availability, changes in routing
- Congestion mechanisms should response promptly to changes but should not over react to transient changes
- Concern to slowly responding mechanisms such TCP congestion window change

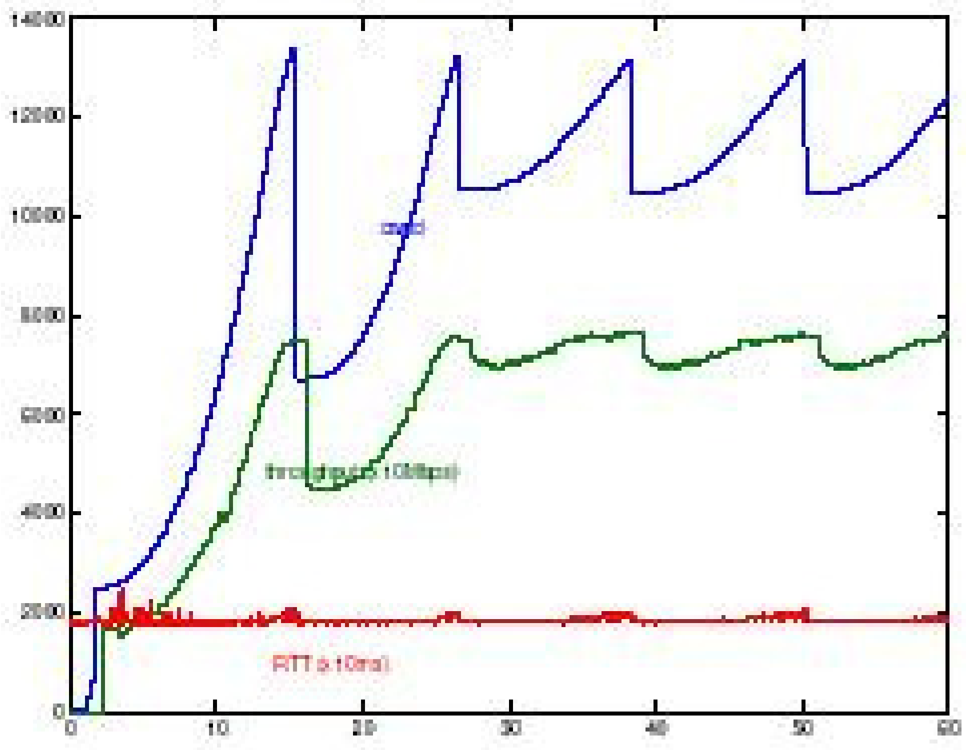
Response Times to Changes

- Responsiviness – number of RTTs of sustained congestion that a (TCP) sender takes to halve its transmission rate
- Agressiveness – maximum increase in the sending rate in one RTT, in packets per second, in the absence of congestion
- Smoothness – Largest reduction in the sending rate during one RTT in a deterministic environment with a periodic packet loss

Oscillation

- Stability
- Predictibility

Stability



Fairness

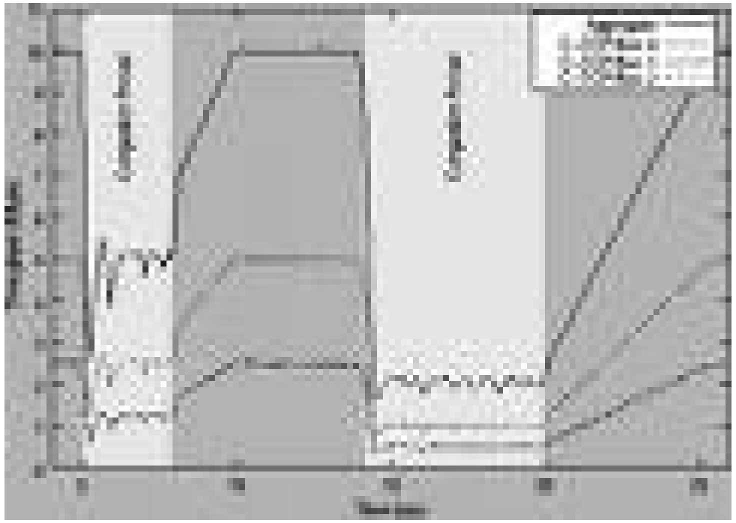
- Among flows of the same protocol and among flows of different protocols
  - Example: fairness between a novel variant of TCP and TCP Reno

- Different indexes of fairness

Fairness

- Usually throughput fairness
- Flows with different RTT
- Flows with different packet size: voice packets and FTP packets

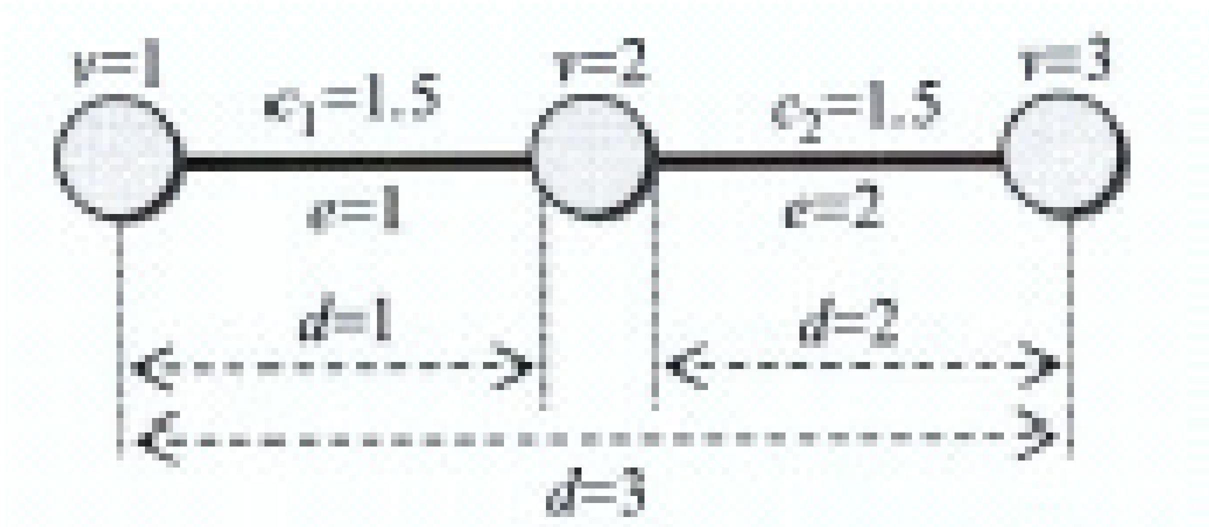
Fairness



Fairness Index

- Max-min fairness:
- A network is fair if each node gets the largest possible share that does not impact nodes with lower share.
- The smallest throughput should be as large as possible.
- Absolute priority to the smallest flows

Max-Min Fairness



Fairness Indexes

- Epsilon-fairness

$$1 - \varepsilon \leq \frac{\min_i x_i}{\max_i x_i}$$

- Proportional fairness – also favours small flows

$$\sum_i \frac{x^*_i - x_i}{x_i}$$

Fairness Indexes

- Jain's Fairness Index

$$\frac{\left[\sum_{i=1}^{n} x_i\right]^2}{n \sum_{i=1}^{n} x_i^2}$$

Fairness

- Convergence to fairness – time taken to converge to a fair situation between existing flow and newly-starting ones
  - Example: TCP variants for high speed networks tend to increase window aggressively and can jeopardize the acquisition of bandwidth by new flows
- Delta fairness – time taken for two flows with same RTT to go from 100/101th and 1/101th shares to $(1+\delta)/2$ and $(1-\delta)/2$ shares of the link
- Number of RTTs to reach epslon fairness

Deployability

- How easy is to deploy a protocol
  - Overhead in packet header size,
  - added complexity to routres and end-systems,
  - complexity of code
  - Failure diagnosis