Low delay Congestion Control

- **Why**
  - Interactive Media (Video, VoIP, RTCweb etc)
  - Delay sensitive data (Safety Monitoring, Financial info, etc)
  - Performance improvements (e.g. Web, Apps etc)
  - Low delay usually means low loss
    - This is significant at lower bit rates as TCP’s rate is approximately inversely proportional to the loss rate => higher loss rate.

- **How**
  - Thread a balance between latency and throughput
    - Minimise path latency
    - Maximise throughput
  - Other issues
    - Packet size
    - Packet pacing
    - Layer 2 cross-layer information
    - Application limited
Metrics: Latency Signals

- Metrics
  - Packet Loss
    - Tells us Queue is full (or AQM has dropped packet)
    - Or Layer 2 losses
    - Loss Differentiation Techniques (e.g. delay trends)
  - Packet Delay
    - RTT/One Way Delay
  - Packet Delay Variance, Jitter, Delay Gradient

- Filtering
  - Minima, Mean, EWMA, Kalman, etc
  - Beware of filter's operation with differing packet rates/sizes
Self-fairness

- Shared knowledge of latency signals with same flows
  - Common measurement of latency signals

- Shared knowledge of latency signals with all flows
  - Have an understanding of metrics relevant to other flow types
TCP/Loss-based-cc competition

- Utilise shared knowledge of latency signals
  - Know when to fall back to delay-based operation once a competing loss-based flow has subsided
- Approaches
  - Modal
    - Can be brittle
  - Functional
    - Finding the appropriate relationship (e.g. CxTCP)
    - Beware of self-inflicted loss
  - Minimum delay is unachievable but on some paths full Queue latency may be ‘in budget’ (e.g. LAN).
A way forward

- Combined metrics
- Appropriate filtering
- Cross-layer info
- Adaptable to the application requirements
  - Video:
    - Variable packet sizes
    - differing importance packets
    - bursting
  - Audio: Smaller packet sizes (non MTU sized)
  - Other…