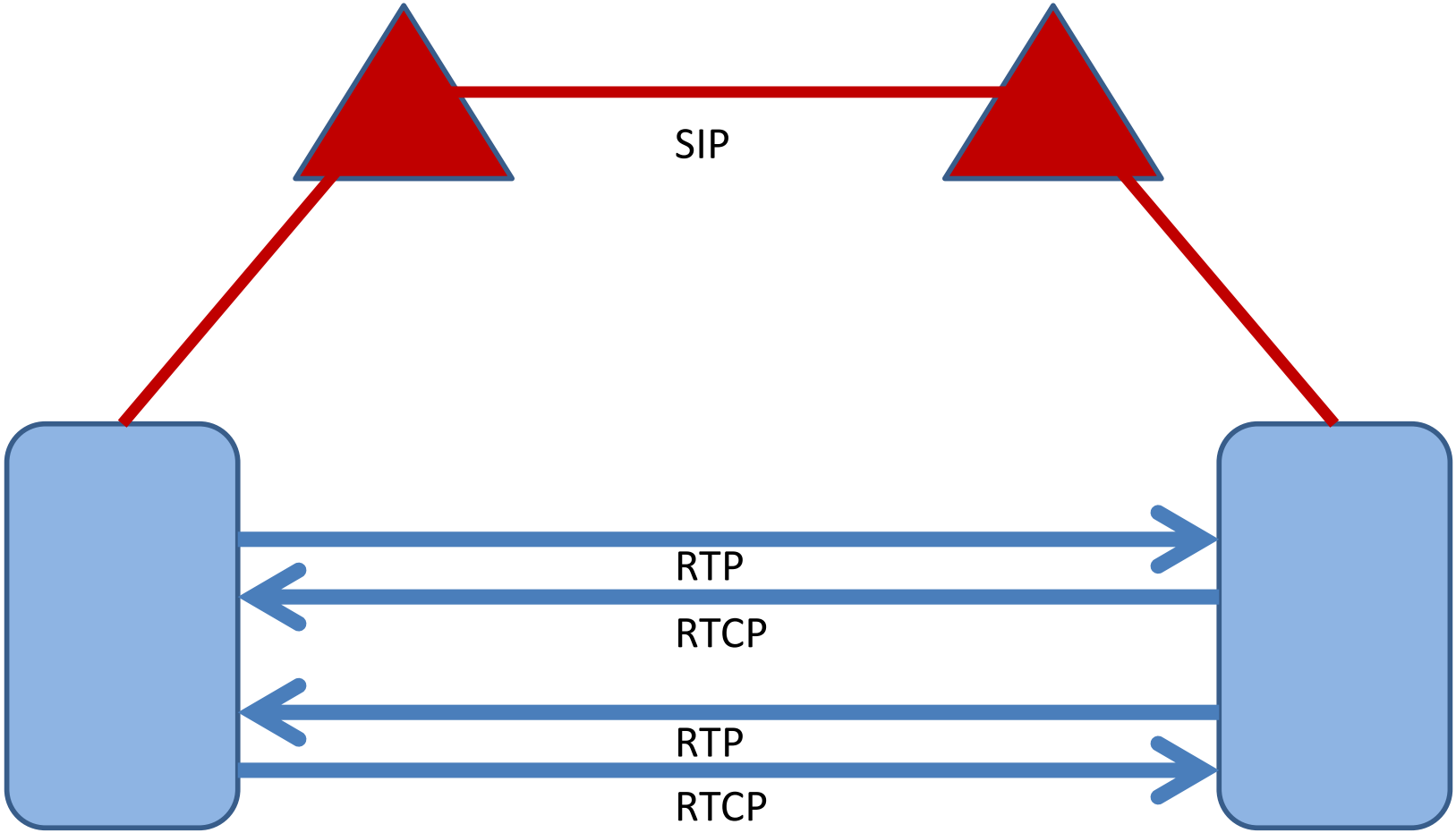
Towards Adaptive Congestion Management for Interactive Real-Time Communications

Dirk Kutscher Mirja Kühlewind Bob Briscoe

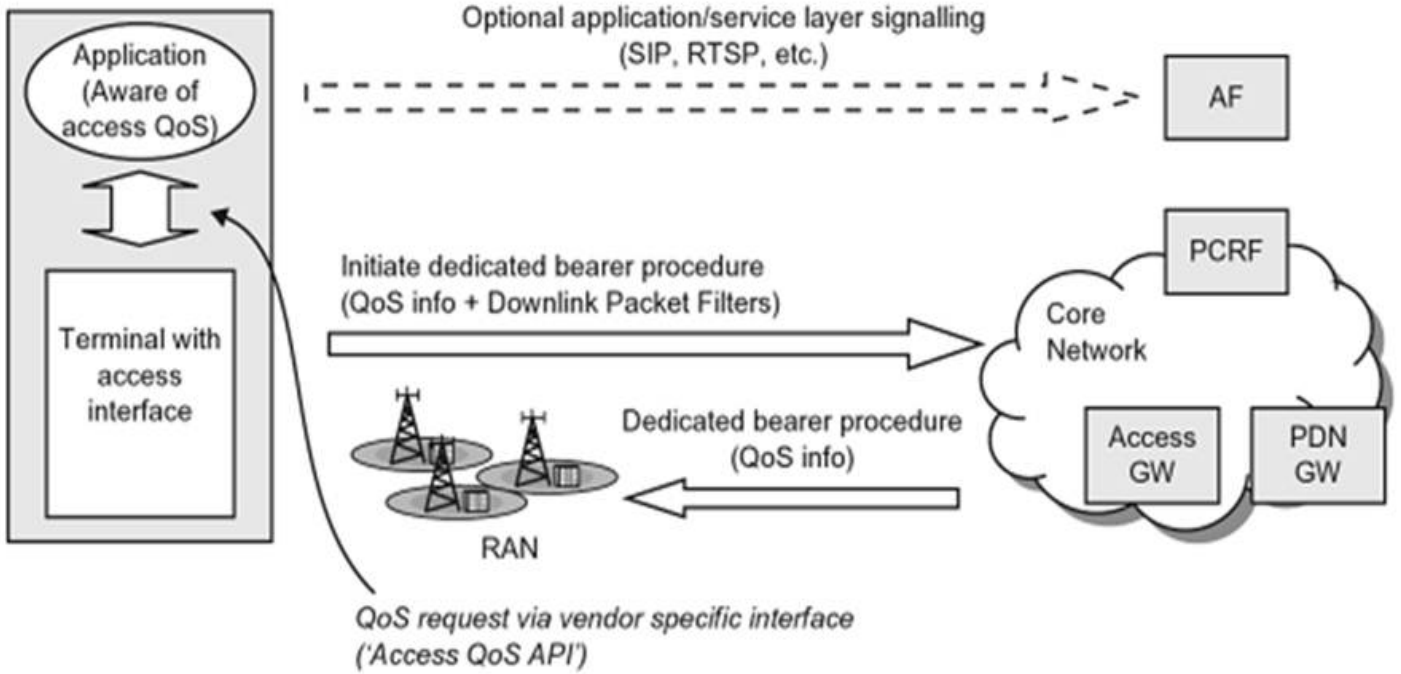
IAB/IRTF Congestion Control Workshop
July 2012

Interactive Real-Time Communication on the Internet



We started here...

...Arrived at Systems



Terminal-initiated QoS control.

- IMS: SIP-based control framework in UMTS and LTE
- Using SIP to trigger QoS reservation
- Dedicated bearers for real-time traffic
- Extensive QoS and control infrastructure (“carrier grade”)

RTCWEB

- Real-time communication without the telco framework
  - No ossified control architecture
  - Web-based model (browser programmability) to foster innovation
  - But obviously no QoS framework
  - Congestion control essentially missing

Essentially: Two Problems

1. Lack of congestion control of IRTC flows
   - Anticipated usage scenarios of RTCWEB to impact TCP and other flows
     - A perceived threat may be enough to trigger operator actions (or to make users abandon it)
   - Uptake of RTCWEB endangered

2. Impact of TCP-based flows on IRTC
   - Poor queue management & loss-based congestion control to induce delays
   - Unhappy RTCWEB users

Challenge

- **Reasonable large-scale IRTC performance over the Internet?**
  - Without having to resort to closed control frameworks?
  - Without needing separate queues for IRTC and TCP?

Way Forward

- Re-think fairness
  - Per flow fairness not helpful
    - Cf. current usage of multi-TCP flows by applications
  - Especially not useful for IRTC
    - E.g., video codecs creating key frames => spikes in sending rates
  - Enable flows to take a bigger share
    - When needed – and when harm to others is acceptable

- Promote, incentivize and ultimately enforce
  - Congestion responsiveness and minimisation
  - For all traffic

Needed Elements

- Better queue management and better congestion signals
  - AQM
- Fast and accurate feedback
  - ECN for RTP (& deploy ECN for TCP too!)
- Adaptiveness on transport and application layers
  - Codec support is there – need profiles and usable APIs

- **Incentives to adapt**
  - **And a suitable fairness model**

Incentives

- Making 'cost' of congestion visible to OS on RTT scales
  - Cost: congestion contribution per flow
- Making 'cost' of congestion visible to network operators

- Example: User A running multiple applications
  - Multiple flows per application session
  - Receiving congestion signals on all of them
  - Only user/application knows what is important
  - Decision on how to re-act should be based on app profiles and user prefs
  - If important video stream cannot be adapted, some TCP flow should (disproportionally)
  - Same model for competing flows of different users / hosts

Congestion Exposure (ConEx)

- Making congestion visible in the network
  - At different time scales, for different purposes
- Senders know their current congestion contribution and declare it to network
- Operator policy can provide necessary incentives to „act responsibly“ and otherwise enforce
- Other usages (on larger time scales)
  - Capacity Provisioning and Traffic engineering according to observed congestion...

Summary

- Want to enable viable IRTC with minimized control plane
- Have to bring down delay and promote adaptiveness
- Many elements already there
- Incentives model based on congestion exposure one of them
- Experiments needed