



Google Congestion Control

Experimental Testbed

Results

Conclusions



Google Congestion Control

Experimental Testbed

Results

Conclusions

- > Enabling real-time communication over the Internet is a hot topic
- > An inter-operable and efficient set of standard protocols is still missing
- > IETF RTC Web
  - > Aims at standardizing a set of protocols to transport real-time flows.
- > W3C Web RTC
  - ➤ Aims at standardizing s set of HTML5 APIs to enable real-time communication within Internet browsers.

- Google Congestion Control
  - Proposed by Google within the RTC Web IETF WG.
  - ➤ Has already been implemented in the Google Chrome and Firefox browsers.
- Experimental investigation of the GCC
  - ➤ Using a controlled testbed which allows bandwidth and propagation times to be set.
  - Investigate to what extent the GCC algorithm is able to
    - 1. Fully utilize the available bandwidth
    - 2. Fairly share the bottleneck bandwidth with concurrent flows
    - 3. Contain queuing delays



**Google Congestion Control** 

Experimental Testbed

Results

Conclusions

# **Google Congestion Control**

- GCC run over the UDP and it encapsulates the audio/video frames in RTP packets
- Applied only to the video streams since the audio streams bitrate are considered negligible

# **Google Congestion Control**

- Figure 1 shows the main components involved in the architecture
- > Two controllers:
  - > a sender-side controller
    - Compute the target sending bitrate As
  - > a receiver-side controller
    - > Compute the rate Ar that is sent to the sender under the conditions

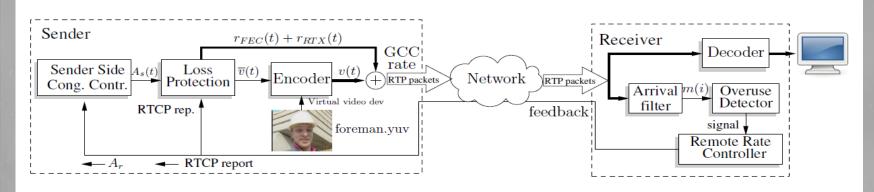


Figure 1: Congestion control architecture

# GCC - The sender-side congestion control

- Loss-based congestion control algorithm
  - > Every time t<sub>k</sub> the k-th RTCP report message arrives at the sender
  - > Every time t<sub>r</sub> the r-th REMB message, which carries Ar, arrives at the sender
- The frequency RTCP reports are sent is time varying
  - ➤ the higher the backward-path available bandwidth, the higher is the RTCP reports frequency.
  - The RTCP reports include the fraction of lost packets f<sub>I</sub>(t<sub>k</sub>)
  - $\rightarrow$  The sender uses  $f_i(t_k)$  to compute the sending rate As( $t_k$ ), measured in kbps

# GCC - The sender-side congestion control

- $\triangleright$  Computation of sending rate As( $t_k$ ):
  - $\triangleright$  X(t<sub>k</sub>): the TCP throughput equation used by the TFRC(TCP-Friendly Rate Control)

$$A_s(t_k) = \begin{cases} \max\{X(t_k), A_s(t_{k-1})(1 - 0.5f_l(t_k))\} & f_l(t_k) > 0.1\\ 1.05(A_s(t_{k-1}) + 1\text{kbps}) & f_l(t_k) < 0.02\\ A_s(t_{k-1}) & \text{otherwise} \end{cases}$$
(1)

➤ When a RMEB is received at time t<sub>r</sub>, As is set as:

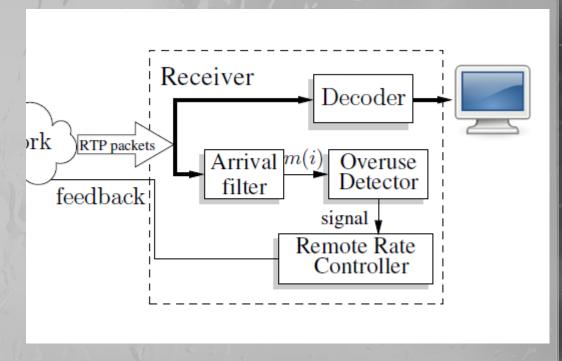
$$A_s(t_r) \leftarrow \min(A_s(t_r), A_r(t_r)).$$

- Delay-based congestion control algorithm
  - Compute Ar according to the following equation:

$$A_r(t_i) = \begin{cases} \eta A_r(t_{i-1}) & \text{Increase} \\ \alpha R(t_i) & \text{Decrease} \\ A(t_{i-1}) & \text{Hold} \end{cases}$$

- t<sub>i</sub>: denote the i-th group of RTP packets carrying a video frame is received
- $> \eta \in [1.005; 1.3], \ \alpha \in [0.8, 0.95]$
- > R(ti): the receiving rate measured in the last 500ms

- Receiver-base controller
  - Arrival-time filter
  - Over-use detector
  - > Remote rate controller



- ➤ Receiver-base controller Arrival-time filter
  - ➤ Goal : estimate the queuing time variation m(t<sub>i</sub>)
  - ➤ Measure the one way delay variation:

$$d_m(t_i)=t_i-t_{i-1}-(T_i-T_{i-1})$$

- > T<sub>i</sub> is the timestamp at which the i-th video frame has been sent
- > t<sub>i</sub> is the timestamp at which the i-th video frame has been received

- > Receiver-base controller Arrival-time filter
  - One way delay variation is considered an the sum of three components[1]:

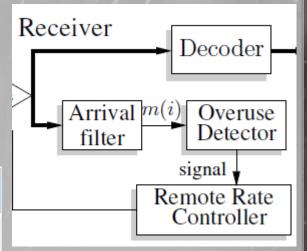
$$d(t_i) = \frac{L(t_i) - L(t_{i-1})}{C(t_i)} + m(t_i) + n(t_i)$$

- $ightharpoonup \frac{L(t_i)-L(t_{i-1})}{C(t_i)}$ : The transmission time variation.
  - L(ti): i-th video frame length; C(ti): estimation of the path capacity
- $\rightarrow$  m(t<sub>i</sub>): the queuing time variation
- > n(t<sub>i</sub>): network jitter which modeled as a Gaussian noise

[1]. H. Lundin et al. Google congestion control algorithm for real-time communication on the world wide web. Draft IETF, 2013.

- Receiver-base controller over-use detector
  - Goal : Produce a signal to remote rate controller based on m(ti)
  - > Rules:

Condition	Signal
if m(ti) increases above a threshold, and keeps increasing for a certain amount of time, or for a certain amount of consecutive frames	Over-use
if m(ti) decreases below a threshold	Under-use
When m(ti) is close to zero	Normal



- > Receiver-base controller- remote rate controller
  - ➤ Goal: compute Ar by using signal produced by the overuse detector
  - > The finite state machine:

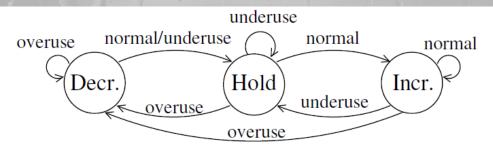
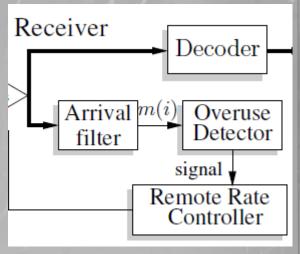


Figure 2: Remote rate controller finite state machine

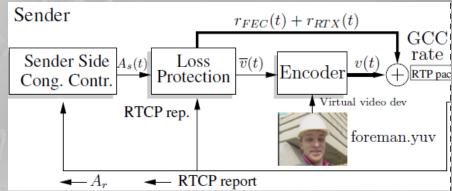


$$A_r(t_i) = \begin{cases} \eta A_r(t_{i-1}) & \text{Increase} \\ \alpha R(t_i) & \text{Decrease} \\ A(t_{i-1}) & \text{Hold} \end{cases}$$

- > Receiver-base controller
  - The signaling from the receiver to the sender is done through REMB message carrying Ar or through RTCP reports.
  - The frequency at which the REMB messages should be sent in the Google Chrome:
    - ➤ If Ar is decreasing, every 1s
    - ➤ If Ar decreases more than 3%, immediately

# GCC – Loss protection

- Forward Error Correction(FEC)
- > Retransmissions to counteract packet losses
- Loss protection is not active
  - $\rightarrow$  v(t) = As(t)
- Loss protection is active
  - $ightharpoonup \overline{v}(t) = As(t) r_{RTX}(t) r_{FEC}(t)$
  - $ightharpoonup r_{FEC}(t)$ :
    - > sending rate of redundant video frames
  - $ightharpoonup r_{RTX}(t)$ 
    - > The retransmission rate
- r<sub>fEC</sub>(t) can not be more than half of the total sending rate A<sub>s</sub>(t)





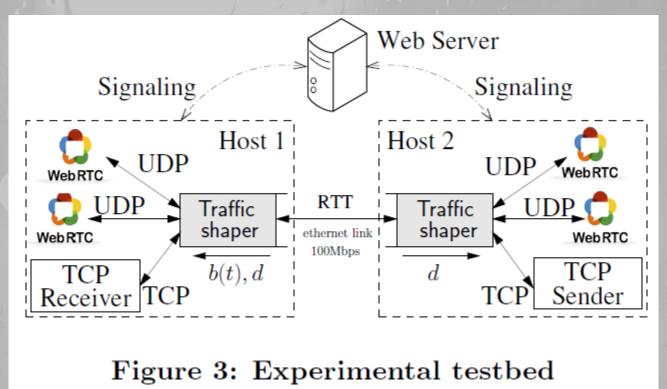
Google Congestion Control

**Experimental Testbed** 

Results

Conclusions

Testbed



- Two host: host1 and host2
- Bottleneck emulating a WAN scenario
  - > Traffic shaper tc<sup>3</sup> to set delays and bandwidth
- > Traffic shaper:
  - ➤ Host 1:
    - > Set a one-way delay d
    - > Set a limitation on the available bandwidth b(t) for the traffic that is received from the Host 2
  - ➤ Host 2:
    - Set the one-way delay d for the traffic that comes from the Host 1
- Minimum round trip time :RTTm = 2d

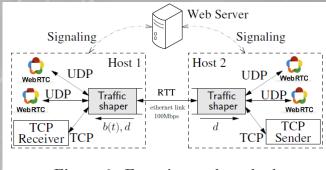


Figure 3: Experimental testbed

- ➤ Buffer size : 60KB
  - > Emulate a typical home gateway
- ➤ Bandwidth range : [500, 2000]kbps
  - > Typical of ADSL uplink speeds
- > Both hosts run a chromium browser which generates the video flows
- ➤ Host 1 is equipped with a TCP receiver
- ➤ Host 2 runs a TCP sender which uses the TCP Cubic congestion control
  - ➤ Log the congestion window, slow-start threshold, RTT, sequence number
- Web server provides the HTML pages

- Metrics employed to assess the performance of GCC in the considered scenarios
  - ➤ Channel Utilization U=R/b :
    - > where b is the known available bandwidth and R is the average received rate
  - ➤ Good Utilization g=v/b :
    - > where v is the average video bitrate without considering the bandwidth used for FEC and retransmissions.
  - Loss radio l=(lost bytes)/(received bytes) :
    - > it is measured by the traffic shaper tool



Google Congestion Control

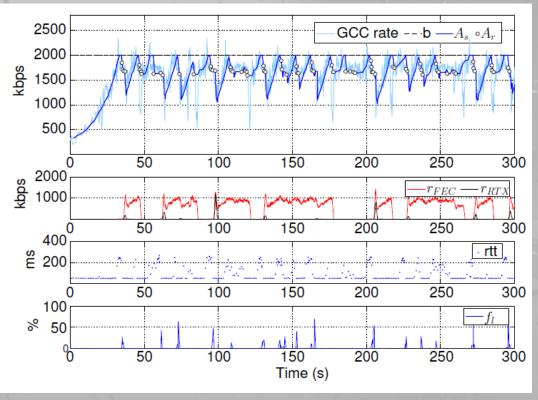
Experimental Testbed

Results

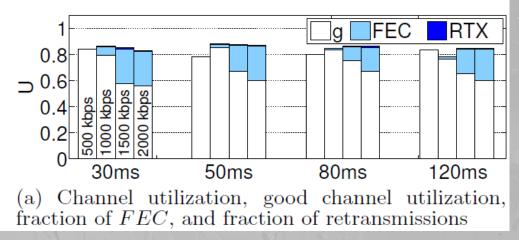
Conclusions

- > One GCC flow over a bottleneck with constant available bandwidth
- ➤ Available bandwidth b(t) ∈{500, 1000, 1500, 2000}kbps
- ➤ Propagation delay RTTm =  $2d \in \{30, 50, 80, 120\}$ ms
- > 16 couples {b(t), RTTm}

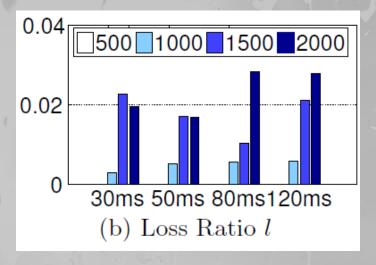
b=2000kbps and RTTm=50ms:



- For each of the considered bottleneck bandwidth bi and round trip delays RTTm
- $\rightarrow$  U = g +  $r_{FEC}(t)/b + r_{RTX}(t)/b$

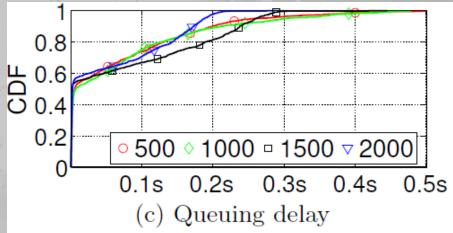


The figure shows that the channel utilization is slightly above 0:8 and is not significantly affected by the RTTm or the available bandwidth.



The figure shows that the loss ratio is not affected by the RTT, but it increases when the bandwidth increases and reaches a maximum value of 0.028.

- ➤ We have computed the queuing delay, defined as Qbi (t) = RTT(t)-RTTm
- > we have computed the four cumulative distribution function (CDF)



The figure shows that the median queuing delay is very close to zero, whereas the 90-th percentile is below 0.25s, confirming that the algorithm is able to contain the queuing delay.

- > One GCC flow over a bottleneck with variable available bandwidth
- Step-like changes of the available bandwidth
- ➤ GCC flow accesses a bottleneck whose bandwidth b(t) changes from a minimum value of 400kbps to 3000kbps after 60s.
- > Conclusion:
  - ➤ GCC reaches a steady-state value of around 2Mbps after a transient time of roughly 30s.

# Let the available bandwidth b( flow adapts the sending rate to

- > Conclusions:
  - b is increased every 100s of 500k

- Figure shows that the sending rate is able to quickly match the variable available bandwidth b(t)
- the measured channel utilization is slightly above 80%.

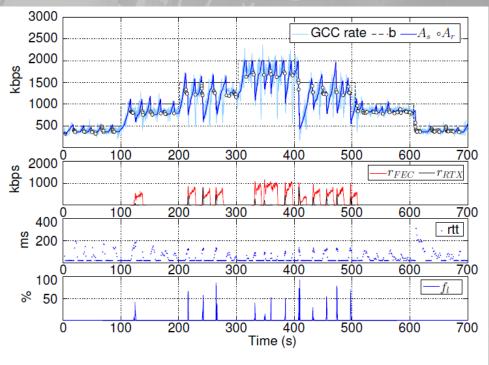


Figure 6: One GCC video flow over a stair-case available bandwidth and  $RTT_m = 50 \text{ms}$ 

- One GCC with one concurrent TCP flow
- ➤ Three different available bandwidths, bi ∈ {1000, 2000, 3000}kbps.
- Propagation delay of RTTm = 50ms.
- > Two cases:
  - 1. A GCC flow a GCC flow starts at t = 0s and a TCP flow enters the bottleneck at t = 100s;
  - 2. a TCP flow starts at t = 0s and a GCC flow is started at t = 100s.

➤ Figure shows the overall channel utilization and the TCP and GCC bandwidth shares obtained in this scenario.

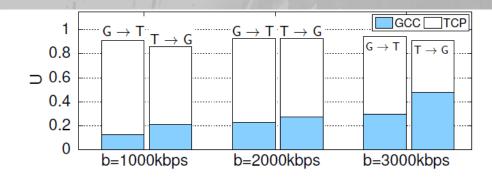


Figure 7: Channel Utilization when one GCC flow (G) shares the bottleneck with a TCP flow (T)

The figure clearly shows that at lower bandwidths GCC is not able to get a fair share.

- Figure shows the dynamics of the GCC flow and a TCP flow when b = 1000kbps and a TCP flow is started before the GCC flow.
- When the GCC flow joins the bottleneck it quickly gets a fair share until when, at around t = 145s
- 2. many consecutive REMB messages are received by the sender and the sending rate is decreased according to the value of Ar

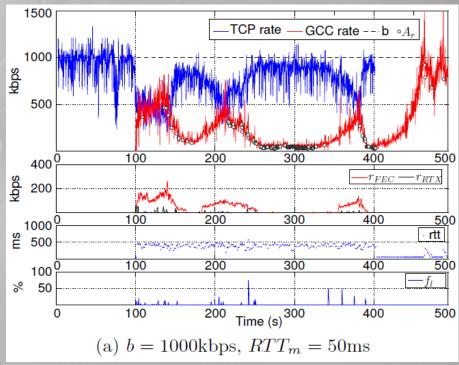
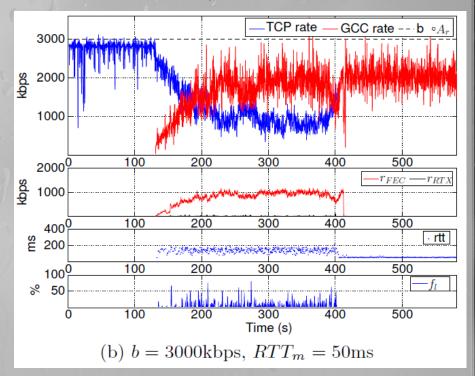


Figure shows the results obtained when b = 3000kbps

- the GCC sending bitrate is driven only by the sender-side controller
- 2. get a larger bandwidth than the TCP flow.



- > Two concurrent GCC flows
- Assess the fairness of two GCC flows when sharing a bottleneck with constant capacity.
- bi∈{1000, 2000, 3000}kbps
- Propagation delay RTTm = 50ms
- For each bi run 5 tests.

It shows that the two GCC flows exhibit an unpredictable behavior

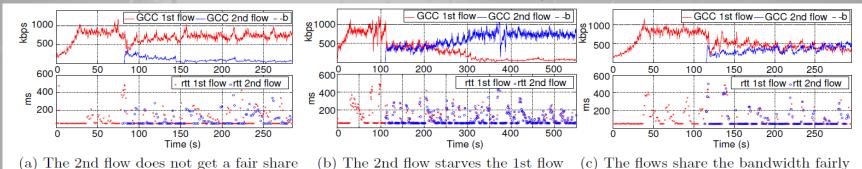


Figure 9: Two GCC flows share a bottleneck with b = 1000 kbps and  $RTT_m = 50 \text{ms}$ 

- a) the second flow obtains only a 10% channel utilization
- b) the first one that only gets 21% of the channel bandwidth
- c) the two ows fairly share the bandwidth.

Figure shows the channel utilization grouped for each of the considered available bandwidths bi.

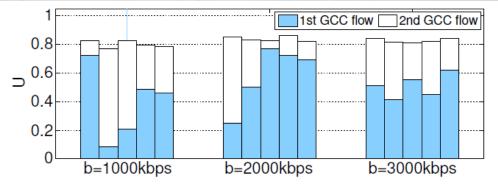


Figure 10: Channel utilization in the case of two GCC flows sharing a bottleneck ( $RTT_m = 50 \text{ms}$ )

- in the case of b = 1000kbps and b = 2000kbps a very poor fairness is obtained
- ➤ b = 3000kbps, the two flow can roughly share the bandwidth fairness. And the first flow always gets a higher bandwidth share.



Google Congestion Control

Experimental Testbed

Results

**Conclusions** 

#### Conclusions

- 1. When a single GCC flow accesses the bottleneck, the channel utilization is above 0.8, even when the available bandwidth changes following a stair-case pattern, and queuing delays are contained
- 2. A video flow controlled by the GCC gets starved when sharing the bottleneck with a TCP flow if the bottleneck capacity is less or equal to 1000 kbps;
- 3. When two GCC video flows share the bottleneck, the algorithm behavior appears unpredictable and exhibit poor fairness.

### **Future Work**

• The cause of the unfairness issues exhibited by GCC when sharing the bottleneck with TCP, or with another GCC flow, is currently under investigation.

