# Enlarge the pre-congestion spectrum usage?

Xavier Marjou
Orange Labs
Lannion, France
xavier.marjou@orange.com

Emile STEPHAN
Orange Labs
Lannion, France
emile.stephan@orange.com

*This position paper is intended for the IAB / IRTF Workshop on Congestion Control for Interactive Real-Time Communication (July 28 2012, Vancouver, Canada). Its content represents early thoughts of the authors and not France Telecom position.*

***Abstract***— This paper proposes use-cases and guidance regarding the congestion control of web-based interactive real-time communications based on the collaboration with all involved entities, including network equipments.

***Keywords; Real Time; Network; Application; Congestion; Control; Exposure; Web;***

## I. INTRODUCTION

With the ongoing standardization of web-based interactive real-time media communications, it is very likely that more and more audio and video RTP/UDP flows will be delivered over the Internet. RFC2914 underlines the crucial need of congestion control in order to prevent the congestion collapse of the Internet and to provide fairness among different competing flows. However, congestion control for RTP/UDP flows has not been deeply investigated yet, which is a new challenge as these applications require low delay and even sometimes high-bandwidth. In this paper, we first present two possible use-cases outlining how real-time interactive communications could well-behave with respect to other Internet applications while still providing good quality (Section II). We thus propose guidance for future work on this topic inside IETF or IRTF (Section III).

## II. USE-CASES

Here are two simplistic use-cases which use congestion control.

### 1) *Use-case A: congestion in an uplink DSL link*

Two web-based applications are in the same browser of a device A. The first application sends a conversational SRTP/UDP flow 1 to device C. The second application sends audiovisual content on a TCP flow 2 to device B. The sum of flow1 and flow2 exceeds the ADSL uplink bandwidth, which creates congestion, as depicted in Figure 1.

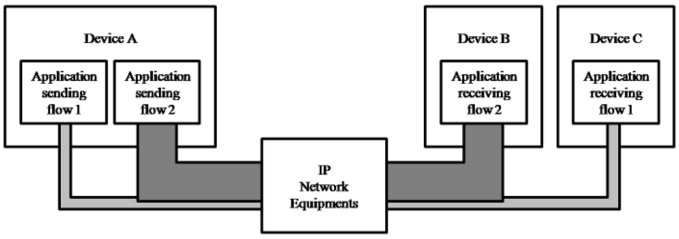


***Figure 1****: congestion in an uplink DSL link*

A congestion control exists, typically with the contributing help of network equipments. Upon congestion or pre-congestion, the application sending flow 2 receives a notification indicating that there is congestion and that it is mainly responsible for the congestion; the application relays this alert to the user, for example with graphical warning hints. The application will have to decrease its bandwidth, maybe with the help of the alerted user; Of course, the user can also stop the application.

Assuming that the user decides to stop the audiovisual content application, the congestion will end, as depicted in Figure 2.

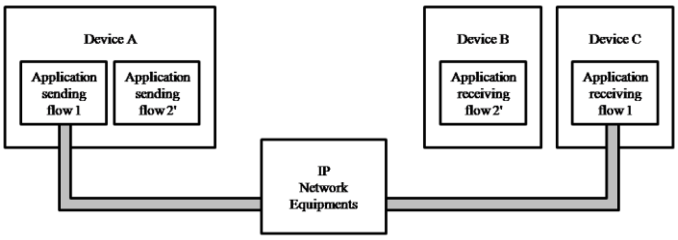


***Figure 2****: congestion finished in the uplink DSL link*

### 2) *Use-case B: congestion in a 3G access-network*

Two devices A and C each host a web-application and are both in the same 3G cell. Device A sends gaming content with a SCTP/DTLS/UDP flow 1 to device B. Device C sends a conversational audio and video SRTP/UDP flow 2 to device D.

The sum of radio resources for flow1 and flow2 exceeds the 3G cell capacity, which creates congestion as depicted in Figure 3.

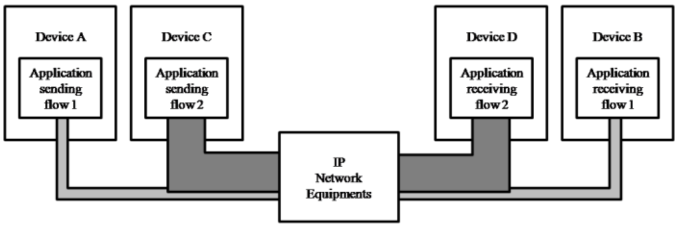


***Figure 3****: congestion in a 3G access-network*

A congestion control exists, typically with the contributing help of network equipments. Upon congestion or pre-congestion, the application sending flow 2 receives a notification indicating that there is congestion and that it is mainly responsible for the congestion; the application relays this alert to the user. The application will have to decrease its bandwidth. The application decides to automatically switch from sending audio and video to sending audio only, the congestion will end, as depicted in Figure 4.

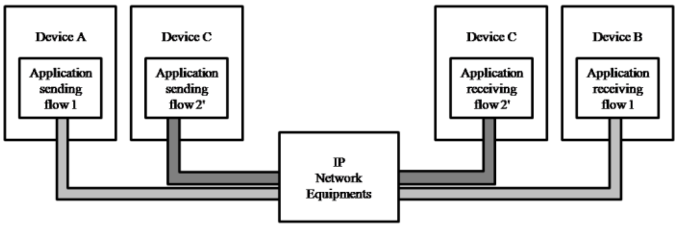


***Figure 4****: congestion finished in the 3G access network*

## III. GUIDANCES

### A. Fairness

As far as the authors know, for now neither the IETF (IPPM WG, AVTCORE WG, AVTEXT WG, etc) nor the IRTF proposed a definition of fairness for resources sharing.

There is a sort of consensus on the usage of fairness for dividing the resource amongst the flows, the applications, the users, etc. Fairness is hard to apply because everyone has its own fairness principles. But what is fairness?

Is it dividing the available bandwidth by the number of sessions? At each node of the path?

Does fairness include the sharing of the available resource of the terminal (CPU, etc)?

What is the right population of interest (sessions, hosts, etc) on which fairness applies?

Can fairness be enforced without sheriff?

Can fairness improve the QoE without tight and real-time indication to and from the path and to and from the application?

IETF or IRTF might work on the definition of Best Current Practice for fairness.

### B. Congestion Avoidance

A transport protocol is designed to enforce the requirements in terms of transport of the data of the application it serves. Per design it is not always network friendly. Furthermore it can not be friendly with the other transport sessions of the other applications and of the other hosts.

UDP does not carry any information related to the QoS experienced. This information, if any, is carried at the application level.

TCP reduces the application throughput when the QoS decreases. Nevertheless its decision relies on the detection of packet losses. Consequently TCP contribute to the creation of situations which are not friendly with Real Time Application.

The solution space encompasses pre-congestion management on a scope wider that one session or one host.

### C. Application Explicit Adaptation Indication

An application reducing its throughput to adapt to a resource limitation might directly inform the path for several reasons:

- To indicate that the path experienced some QoS issue;
- To indicate that the freed resources should be used cautiously. This should enforce the effort of adaptation made by the application because the bandwidth freed will decrease the risk of pre-congestion on the path for a longer period of time.

The intent of this paper is not to provide solutions. Nevertheless this might be performed using RE-ECN like solution.

## IV. CONCLUSION

In this paper, we have shown 2 points:

- There is a need for a global pre-congestion framework working for all main protocols used in web-based applications.
- The user of the application, the application of the content provider and the network equipments along the data path must be able to evaluate the contribution of a flow to the overall congestion.

Regarding pre-congestion exposure, the work done in the IETF CONEX working group already tackles the work for TCP. We suggest discussing 2 proposals during the workshop:

- Extending such work to any other transport (UDP, SCTP, etc)
- Extending sockets APIs to communicate pre-congestion information between the application, the transport level and the IP level.

.