Pre-congestion notification in mobile networks
draft-jobert-tsvwg-pre-congest-notif-mobile-00.txt

Status of this Memo

   This Internet-Draft is submitted in full conformance with the
   provisions of BCP 78 and BCP 79.

   This document may contain material from IETF Documents or IETF
   Contributions published or made publicly available before November
   10, 2008. The person(s) controlling the copyright in some of this
   material may not have granted the IETF Trust the right to allow
   modifications of such material outside the IETF Standards Process.
   Without obtaining an adequate license from the person(s) controlling
   the copyright in such materials, this document may not be modified
   outside the IETF Standards Process, and derivative works of it may
   not be created outside the IETF Standards Process, except to format
   it for publication as an RFC or to translate it into languages other
   than English.

   Internet-Drafts are working documents of the Internet Engineering
   Task Force (IETF), its areas, and its working groups. Note that
   other groups may also distribute working documents as Internet-
   Drafts.

   Internet-Drafts are draft documents valid for a maximum of six
   months and may be updated, replaced, or obsoleted by other documents
   at any time. It is inappropriate to use Internet-Drafts as
   reference material or to cite them other than as "work in progress."

   The list of current Internet-Drafts can be accessed at
   http://www.ietf.org/ietf/1id-abstracts.txt

   The list of Internet-Draft Shadow Directories can be accessed at
   http://www.ietf.org/shadow.html

   This Internet-Draft will expire on September 5, 2012.

Copyright Notice

Abstract

   Mobile networks may be divided into two main segments: the radio
   segment, and the wireline segment. This document highlights that the
   algorithms leading to pre-congestion notification (e.g. ECN marking)
   are usually significantly different for these two segments, and not
   defined or documented in general over the radio segment. It also
   explains that using ECN bits leads to having a unique signal for
   notifying a pre-congestion related to two separate segments with
   very different notification algorithms. Some consequences are
   questioned, as well as the potential benefits of being able to
   identify where the congestion comes from. This document finally
   discusses the possibility to take into account the radio conditions
   of the terminals when counting the volume of congestion over the
   radio segment.

Table of Contents

1. Introduction

   Mobile networks may be divided into two main segments, very unlike
   by nature: the radio segment, and the wireline segment. These two
   portions of a mobile network may experience QoS degradation due to
   excess traffic. Therefore, being able to notify about a so-called
   pre-congestion (e.g. using ECN marking) can be considered as a
   useful feature for these two segments.

   This document highlights that the algorithms leading to pre-
   congestion notification (e.g. ECN marking) are usually significantly
   different for these two segments. In particular, they are in general
   more complex on the radio segment, and not really defined or
   documented. Depending on the intended purpose, different algorithms
   might be designed over this segment and it is therefore important
   that they are understood and documented somewhere before being
   applied to a specific scenario.

   This document also reminds the typical IP layers in presence in
   mobile networks, e.g. due to the use of GTP tunnels. It highlights
   that the standardized ECN coding in the header of IP packets leads
   to having a unique signal for communicating to the receiver of the
   flows pre-congestion information potentially related to two separate
   segments with very different notification algorithms. The document
   suggests that the consequences of a common interpretation of this
   unique signal need to be assessed more in details and raises the
   question of potential benefits in being able to identify where the
   congestion comes from (e.g. using separated signals to inform about
   pre-congestion over these two segments).

   Finally, this document discusses the use of pre-congestion
   notification over the radio segment in some use cases related to the
   IETF ConEx WG, e.g. where the volume of congestion is counted. It
   advocates that counting the number of bytes transmitted over the
   radio segment during a pre-congestion period may not be the best
   approach to provide incentive to reduce network usage during these
   periods, because the terminals in bad radio conditions require more
   radio resources compared to the terminals in good radio conditions
   to reach the same rate. The possibility to take into account the
   radio conditions of the terminals when counting the volume of
   congestion over the radio segment is briefly introduced.

2. Conventions used in this document

   The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
   "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
   document are to be interpreted as described in RFC-2119 [RFC2119].

   In this document, these words will appear with that interpretation
   only when in ALL CAPS. Lower case uses of these words are not to be
   interpreted as carrying RFC-2119 significance.

   IP E2E: the end-to-end IP layer related to the end application.

   IP TNL: the transport IP layer supporting the GTP tunnel in mobile
   networks.

   UE: User Equipment

   NB: NodeB

   ECN: Early Congestion Notification

   CQI: Channel Quality Indicator

   AQM: Active Queue Management

   RED: Random Early Detection

3. Radio and wireline segments in mobile networks

   Mobile networks may be divided into two main segments, very
   different by nature: the radio segment, and the wireline segment.
   The figure 1 below illustrates these two segments with the example
   of an LTE/EPC network, and also shows the two IP layers in presence
   in the backhaul and core portions:

   o IP E2E layer: the end-to-end IP layer related to the end
      application

   o IP TNL layer: the transport IP layer which supports the GTP
      tunnel

```
+--------------------------------------------------------------------+
|                                                                    |
|     .-----.           .------.          .----.          .----.     |
|    /       \         /        \        /      \        /      \    |
|eUE-  Radio  --eNB- Backhaul -S-GW-  Core  -PDN-GW-Internet-Server| |
|    \       /         \        /        \      /        \      /    |
|     '-----'           '------'          '----'          '----'     |
|                                                                    |
|.-----.                                                      .-----. |
||Appli| --------------------------------------------------- |Appli| |
||-----|                                                      |-----| |
||Sess | --------------------------------------------------- |Sess | |
||-----|                                         .-----.----. |-----| |
|| IP  | ------------------------------------- | IP | IP | - | IP  | |
||-----|   .------.-----.   .----.-----.   |----+----|   |-----| |
||     |   |      | GTP | - | GTP | GTP | - | GTP |   |   |     | |
||     |   |      |-----|   |----+----|   |-----|   |   |     | |
|| L2  | - | L2   | UDP | - | UDP | UDP | - | UDP |   |   |     | |
||radio|   | radio|-----|   |----+----|   |-----| L2 | - | L2  | |
||     |   |      | IP  | - | IP | IP  | - | IP  |   |   |     | |
||     |   |      |-----|   |----+----|   |-----|   |   |     | |
||     |   |      | L2  | - | L2  | L2  | - | L2  |   |   |     | |
||-----|   |------+-----|   |----+-----|   |-----+----|   |-----| |
|| L1  | - | L1  | L1  | - | L1  | L1  | - | L1  | L1 |   | L1  | |
||'-----'   '------'-----'   '-----'-----'   '-----'----'   '-----' |
|                                                                    |
| '-------v-------' '----------------------v-------------------'     |
|   Radio segment                 Wireline segment                   |
|                                                                    |
+--------------------------------------------------------------------+
```
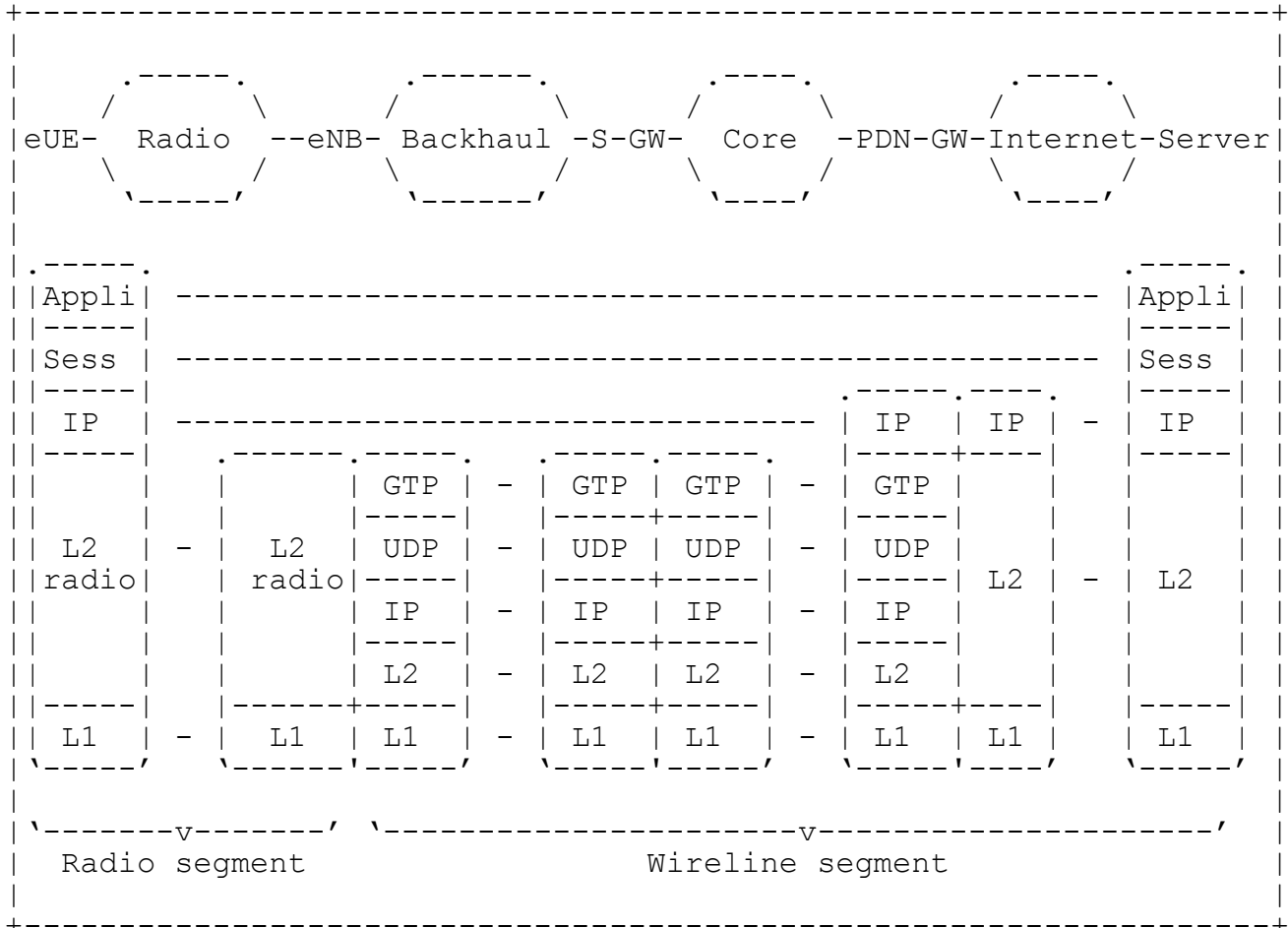
       Figure 1 – Radio and wireline segments in mobile networks (LTE/EPC
                        example, data plane)

4. Pre-congestion notification in radio and wireline segments

   The two portions of a mobile network shown in figure 1 may
   experience QoS degradation due to excess traffic. Therefore, being
   able to notify about a pre-congestion (e.g. using ECN marking) can
   be considered as a useful feature for these two segments.

   It is important to stress that the algorithms leading to pre-
   congestion notification (e.g. ECN marking) are usually significantly

different for these two segments. In particular, they are in general
more complex on the radio segment, and not really defined or
documented. Depending on the intended purpose, different algorithms
might be designed over this segment and it is therefore important
that they are understood and documented somewhere before being
applied to a specific scenario.

4.1. Pre-congestion notification in wireline segment

The algorithms for pre-congestion notification in an IP mobile
wireline network are pretty well documented in IETF documents.
Indeed, it is expected to correspond to classical ECN marking
algorithms in IP routers.

[RFC3168] ("The Addition of Explicit Congestion Notification (ECN)
to IP") depicts the general ideas of ECN marking, which are based on
Active Queue Management (AQM) and for example Random Early Detection
(RED) principles (e.g. [RFC 2309], "Recommendations on Queue
Management and Congestion Avoidance in the Internet").

In particular, [RFC3168] mentions the following:

o "AQM drops packets based on the average queue length exceeding a
   threshold, rather than only when the queue overflows."

o "AQM can set a Congestion Experienced (CE) codepoint in the
   packet header instead of dropping the packet, when such a field
   is provided in the IP header and understood by the transport
   protocol."

o "For a router, the CE codepoint of an ECN-Capable packet SHOULD
   only be set if the router would otherwise have dropped the packet
   as an indication of congestion to the end nodes."

o "We expect that routers will set the CE codepoint in response to
   incipient congestion as indicated by the average queue size,
   using the RED algorithms suggested in [FJ93, RFC2309]. To the
   best of our knowledge, this is the only proposal currently under
   discussion in the IETF for routers to drop packets proactively,
   before the buffer overflows. However, this document does not
   attempt to specify a particular mechanism for active queue
   management, leaving that endeavor, if needed, to other areas of
   the IETF."

These mechanisms aim mainly at interacting with TCP, in order to
reduce the bandwidth consumption when pre-congestion is detected.

4.2. Pre-congestion notification in radio segment

The algorithms for pre-congestion notification in the radio segment
of a mobile network are however expected to be significantly
different from the wireline segment and more complex in general.

Indeed, in the radio segment, a limited number of radio resources
are available for all User Equipment (UE) of a cell. Every TTI
(Transmission Time Interval), radio resources are dynamically
allocated by the radio scheduler to the active UEs of the cell.

In addition to the radio resources allocation, the quality of the
radio channel of a given UE is a key parameter determining the
achievable throughput for this UE: indeed, more complex and
efficient radio Modulation and Coding Schemes (MCS) can be used when
the UE is in very good radio conditions, leading to a higher
throughput per radio resource. On the contrary, for the same amount
of radio resources allocated, a UE in poor radio reception
conditions requires more robust and simpler MCS and will experience
lower bit rates than the same UE in good radio conditions. Hence, a
UE in poor radio conditions will need much more radio resources to
reach the same throughput compared to a UE in good radio conditions.

The radio conditions are measured over the downlink channel by the
UE and are reported to the mobile network, by means of Channel
Quality Indicator (CQI) values.

An algorithm aiming at notifying pre-congestion in the radio segment
is therefore expected to take into account different parameters, as
for instance:

o Average queue length exceeding a threshold, as for RED

o Amount of radio resources used by a particular UE and/or by all
   the active UEs of the cell

o Radio conditions of a particular UE

The details of the algorithm for ECN bits marking in the radio
segment, when supported, can however be considered as mainly
proprietary at the moment, and is not known in general by the
network operator.

On this aspect, [5] (draft "Mobile Communication Congestion Exposure
Scenario") presented in the IETF ConEx WG mentions the following:

o "ECN is already partially introduced into 3GPP networks: Section
   11.6 in [3GPP.36.300] specifies the usage of ECN for congestion
   notification on the radio link (between eNB and UE), and
   [3GPP.26.114] specifies how this can be leveraged for voice codec
   adaptation."

However, when looking at these 3GPP documents, they give only
general information about the expected behavior at the UE (e.g.
codec rate reduction), but not about the details of the ECN marking
mechanism by the NB. For instance, the section 11.6 of [3] (TS
36.300) dealing with Explicit Congestion Notification is copied
below:

o "The eNB and the UE support of the Explicit Congestion
   Notification (ECN) is specified in Section 5 of [35] (i.e., the
   normative part of [35] that applies to the end-to-end flow of IP
   packets), and below. This enables the eNB to control the initial
   codec rate selection and/or to trigger a codec rate reduction.
   Thereby the eNB can increase capacity (e.g., in terms of number
   of accepted VoIP calls), and improve coverage (e.g. for high bit
   rate video sessions)."

o "The eNB should set the Congestion Experienced (CE) codepoint
   ('11') in PDCP SDUs in the downlink direction to indicate
   downlink (radio) congestion if those PDCP SDUs have one of the
   two ECN-Capable Transport (ECT) codepoints set. The eNB should
   set the Congestion Experienced (CE) codepoint ('11') in PDCP SDUs
   in the uplink direction to indicate uplink (radio) congestion if
   those PDCP SDUs have one of the two ECN-Capable Transport (ECT)
   codepoints set."

Some algorithms for marking ECN bits in the radio segment can be
found in the literature; they may take into account some of the
parameters listed before. These algorithms may aim also at
interacting with TCP, in order to improve either the overall
capacity or the fairness between UEs, but other design choices are
also possible.

Anyway, there is no well-identified default ECN marking algorithm for the radio segment, so, it can be assumed that an actual implementation will make certain design choices and favor certain criteria compared to others (e.g. capacity vs fairness).

Those choices are not necessarily expected to meet the objectives of a congestion-volume counting approach like discussed in the IETF ConEx WG, especially when the exact algorithm is not known by the network operator. They might not be the most optimal choices either.

4.3. General remarks about pre-congestion notification in mobile networks

It has been shown that the criteria considered for pre-congestion notification in the wireline segment and in the radio segment are significantly different.

It has also been shown also that different algorithm designs are possible for the pre-congestion notification over the radio segment, and that the details are not necessarily known, which makes the mechanism difficult to be applied to specific scenarios.

There might be benefits in particular in providing more details about pre-congestion notification in the radio segment, in order to better understand the scenario on which it is suitable.

As it will be explained in the next section, there might be also benefits in being able to identify where the congestion happened (radio vs wireline segment), in order to potentially take different actions.

5. ECN bits in IP E2E layer: a single signal to carry pre-congestion notification related to two separate segments

The figure 1 presented before reminds the typical IP layers in presence in mobile networks (IP E2E and IP TNL), e.g. due to the use of GTP tunnels.

This figure highlights that the standardized ECN coding in the header of IP E2E layer packets leads to having a unique signal for communicating to the receiver of the flows pre-congestion information potentially related to two separate segments, with very

different notification algorithms. Hence, pre-congestion located in
the radio segment cannot be distinguished from those of the wireline
segment.

As examples, the following cases may happen when pre-congestion
occurs in the downstream direction (i.e. from the server to the UE).
The behavior described in [RFC6040] ("Tunnelling of Explicit
Congestion Notification") consisting in propagating a 'CE' marking
at the output of the IP tunnel is assumed.

o Pre-congestion happened in the radio segment only: the ECN bits
   of the IP E2E layer are expected to be marked to 'CE' by the NB
   as to notify the UE about this radio pre-congestion.

o Pre-congestion happened in the wireline segment only: the ECN
   bits of the IP TNL layer are expected to be marked to 'CE' by an
   IP router in the wireline network. The NB is then expected to
   propagate this 'CE' marking to the ECN bits of the IP E2E layer
   as to notify the UE about this wireline pre-congestion.

o Pre-congestion happened in the radio segment and in the wireline
   segment simultaneously: the ECN bits of the IP TNL layer are
   expected to be marked to 'CE' by an IP router in the wireline
   network. The ECN bits of the IP E2E layer are also expected to be
   marked to 'CE' by the NB as to notify the UE about this radio
   pre-congestion.

These illustrative examples show that the UE may not know where the
pre-congestion comes from when using the ECN bits of the IP E2E
layer.

One might argue that the UE is not interested in knowing the
location of the pre-congestion, and that the ECN marking is an end-
to-end mechanism. This is correct under the assumption that the ECN
marking criteria are consistent over the entire network. In the
examples of a mobile network provided before, it is not the case.

Therefore, an ECN marking could be misinterpreted by the UE (e.g. as
a wireline pre-congestion instead of a radio pre-congestion, or
vice-versa), with potentially inappropriate actions. A wireline pre-
congestion notification might also be "erased" by a radio pre-
congestion notification.

It is important to remind, as explained before, that the algorithms
leading to ECN marking are significantly different for these two
segments. The consequences of a common interpretation of this unique

signal corresponding to the ECN marking in the IP E2E layer need therefore to be assessed more in details, and probably depends on the actions that are triggered (e.g. interaction with TCP, congestion-volume counting…).

Based on this discussion, this document raises the question of potential benefits in being able to identify where the congestion comes from (e.g. using separated signals to inform about pre-congestion over these two segments).

The main question is in particular to determine whether the ECN bits of the IP E2E layer correspond to the best signal for indicating a pre-congestion happening in the radio segment.

Indeed, the actions to be taken when a pre-congestion occurred may depend on the location of the pre-congestion (ECN marking on the radio could lead to less strict actions compared to a backhaul congestion, depending on the details of the ECN marking algorithm over the radio segment). Moreover, although not the main purpose of the mechanism, it might be useful also for the network management to identify where the congestions are located.

These arguments might be further developed in a future revision of this paper, together with potential proposals to define separate signals to indicate pre-congestion notification over these two different segments of a mobile network. This type of approach would be compared to the complexity of defining separate signals.

6. Options for congestion-volume counting over the radio segment

The current proposal in the IETF ConEx WG for counting congestion-volume is as follows: "the "congestion volume" is defined to be the total number of bytes marked as congested" (see [6] - draft "Congestion Exposure (ConEx) Concepts and Abstract Mechanism").

As it has been explained in the paragraph 4.2, the radio conditions of a UE are an important parameter which determines the amount of radio resources required to reach a given throughput. Indeed, a UE in poor radio conditions will need much more radio resources to reach the same throughput compared to a UE in good radio conditions.

For this reason, when applying use cases related to ConEx where the volume of congestion is counted, it could be of interest to take into account the radio conditions of the UE.

For example, each byte marked as pre-congested may be weighted by a multiplicative factor depending on the UE radio conditions instead of simply counting the number of bytes transmitted over the radio segment during a pre-congestion period.

N thresholds (e.g. corresponding to different ranges of CQI values reported by the UE) might for instance be defined to refine the way the data volume is counted during pre-congestion periods. For instance, if N=3:

o Excellent radio conditions: the congestion-volume is counted with a factor F=1, i.e. the number of bytes transmitted over the radio segment during a congestion period is counted

o Degraded radio conditions: the congestion-volume is weighted with a factor F=2, i.e. twice the number of bytes transmitted over the radio segment during a congestion period is counted

o Bad radio conditions: the congestion-volume is weighted with a factor F=5, i.e. five times the number of bytes transmitted over the radio segment during a congestion period is counted

Another alternative would be that the pre-congestion notification probability (e.g. ECN marking probability) would take into account the radio conditions of the UE. Basically, the packets of a UE in bad radio conditions would be marked more often under pre-congestion periods than those of a UE in good radio conditions. This would provide an equivalent mechanism as the multiplicative factor described above.

This type of mechanism would provide incentive to end users in bad radio conditions to delay their non-urgent network consumptions. Of course, the count would be only active when the cell is considered as in pre-congestion (according to criteria to be further defined).

This proposal might be further developed in a future revision of this paper.

7. Conclusions

This document has reminded that mobile networks may be divided into two main segments, very different by nature: the radio segment, and the wireline segment. It also highlighted that the algorithms leading to pre-congestion notification (e.g. ECN marking) are

usually significantly different for these two segments, and not
defined or documented in general over the radio segment. It also
explained that using ECN bits leads to having a unique signal for
notifying a pre-congestion related to two separate segments with
very different notification algorithms. Some consequences of a
common interpretation of this unique signal have been questioned, as
well as the potential benefits in being able of identifying where
the congestion comes. This document finally discussed the
possibility to take into account the radio conditions of the
terminals when counting the volume of congestion over the radio
segment.

This document proposes that:

o   The main families of algorithms leading to pre-congestion
    notification (e.g. ECN marking) in the radio segment would be
    documented somewhere. This might involve Standard Development
    Organisms beyond the IETF. It is however considered useful
    information for IETF work.

o   The signal indicating a pre-congestion over the radio segment
    would be discussed. The main question is in particular to
    determine whether the ECN bits of the IP E2E layer correspond to
    the best signal or if a separate signal should be defined.

o   For the use cases where congestion-volume are counted (as
    discussed in the IETF ConEx WG), the radio conditions of the UE
    are taken into account in the count, either via the introduction
    of a multiplicative factor or with a pre-congestion notification
    probability (e.g. ECN marking probability) taking into account
    the radio conditions.

Some proposals contained in this document might be further developed
in a future revision of this paper.

8. Security Considerations

   <Add any security considerations>

9. IANA Considerations

   <Add any IANA considerations>

10. References

10.1. Normative References

[1]   Bradner, S., "Key words for use in RFCs to Indicate
        Requirement Levels", BCP 14, RFC 2119, March 1997.

[2]   Crocker, D. and Overell, P.(Editors), "Augmented BNF for
        Syntax Specifications: ABNF", RFC 2234, Internet Mail
        Consortium and Demon Internet Ltd., November 1997.

[3]   3GPP TS 36.300 V11.0.0 "3rd Generation Partnership Project;
        Technical Specification Group Radio Access Network; Evolved
        Universal Terrestrial Radio Access (E-UTRA) and Evolved
        Universal Terrestrial Radio Access Network (E-UTRAN); Overall
        description; Stage 2 (Release 11)", December 2011.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
          Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2234] Crocker, D. and Overell, P.(Editors), "Augmented BNF for
          Syntax Specifications: ABNF", RFC 2234, Internet Mail
          Consortium and Demon Internet Ltd., November 1997.

[RFC3168] Ramakrishnan, Floyd, Black "The Addition of Explicit
          Congestion Notification (ECN) to IP", RFC 3168, September
          2001.

[RFC2309] Braden et al "Recommendations on Queue Management and
          Congestion Avoidance in the Internet", RFC 2309, April
          1998.

[RFC6040] Briscoe "Tunnelling of Explicit Congestion Notification",
          RFC 6040, November 2010.

10.2. Informative References

[4]   Faber, T., Touch, J. and W. Yue, "The TIME-WAIT state in TCP
        and Its Effect on Busy Servers", Proc. Infocom 1999 pp. 1573-
        1583.

[5]   Kutscher, Winter, Krishnan, Zhang "Mobile Communication
        Congestion Exposure Scenario", October 2011.

   [6]   Mathis, Briscoe "Congestion Exposure (ConEx) Concepts and
          Abstract Mechanism", October 2011.

   [Fab1999] Faber, T., Touch, J. and W. Yue, "The TIME-WAIT state in
          TCP and Its Effect on Busy Servers", Proc. Infocom 1999
          pp. 1573-1583.

11. Acknowledgments

   This document was prepared using 2-Word-v2.0.template.dot.

Authors' Addresses

   Sébastien Jobert
   France Telecom Orange
   2 avenue Pierre Marzin
   22300 LANNION, FRANCE

   Email: sebastien.jobert@orange.com

   Isabelle Hamchaoui
   France Telecom Orange
   2 avenue Pierre Marzin
   22300 LANNION, FRANCE

   Email: isabelle.hamchaoui@orange.com