Computer Networks 55 (2011) 2380–2391

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

ECN verbose mode: A statistical method for network path congestion estimation ☆

Rémi Diana [a,*], Emmanuel Lochin [b,c]

[a] TéSA/CNES/Thales, Toulouse, France
[b] CNRS, LAAS, 7 avenue du Colonel Roche, F-31077 Toulouse, France
[c] Université de Toulouse, UPS, INSA, INP, ISAE, LAAS, F-31077 Toulouse, France

ARTICLE INFO

Article history:
Received 9 April 2010
Received in revised form 13 October 2010
Accepted 2 April 2011
Available online 9 April 2011
Responsible Editor: B. Li

Keywords:
Congestion estimation
ECN
Measurements

ABSTRACT

This article introduces a simple and effective methodology to determine the level of congestion in a network with an ECN-like marking scheme. The purpose of the ECN bit is to notify TCP sources of an imminent congestion in order to react before losses occur. However, ECN is a binary indicator which does not reflect the congestion level (i.e. the percentage of queued packets) of the bottleneck, thus preventing any adapted reaction. In this study, we use a counter in place of the traditional ECN marking scheme to assess the number of times a packet has crossed a congested router. Thanks to this simple counter, we drive a statistical analysis to accurately estimate the congestion level of each router on a network path. We detail in this paper an analytical method validated by simulations which demonstrate the feasibility and the accuracy of the concept proposed and illustrate its use in a realistic scenario. We conclude this paper with possible applications and expected future work.

1. Introduction

While dropping packets to prevent congestion was considered as a paradox, many studies have shown the undeniable assets of the Explicit Congestion Notification flag [13]. The story starts in 1994 when Sally Floyd shows that this notification allows to increase TCP performances [3] and later in [9], where the authors reach similar conclusion concerning the web traffic. At last, Aleksandar Kuzmanovic in "The Power of Explicit Congestion Notification" [8] investigates the pertinence of ECN and demonstrates once again, that ECN's users will obtain better performances even if all the Internet is not fully ECN-capable.

The following study [10] published in 2004 precises that ECN is only used by 2.1% of computers and that this low percentage can be partly explained by firewall, NAT and other *middle-boxes* of the Internet which reset (without any justification) the ECN flag. However, this is definitely not the main reason. Indeed, although this flag is currently implemented both in end-hosts (GNU/Linux, Mac OSX and Windows Vista) and inside the core network (Cisco IOS implements a RED/ECN variant called WRED/ECN), ECN remains surprisingly disabled by default for all these systems. Concerning end-hosts, this might appear paradoxical. While today CUBIC and Compound TCP variants are enabled by default (respectively in GNU/Linux and Windows Vista) and are still under debate concerning their friendliness with the current Newreno TCP version, a proved mechanism as ECN is not.

We believe this trend has two main reasons: firstly, this is partly due to the behaviour of TCP face to ECN marked packets. Indeed, the goal of the ECN bit is to notify TCP sources of an imminent congestion but this binary

☆ Part of these results has been presented at IEEE Infocom 2010 Work in Progress track.

* Corresponding author. Address: ISAE, Campus ENSICA,1 place Emile Blouin, BP 75064, 31033 TOULOUSE Cedex 5, France.

E-mail addresses: remi.diana@isae.fr (R. Diana), emmanuel.lochin@isae.fr (E. Lochin).

1389-1286/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.comnet.2011.04.001

indicator does not reflect the real network congestion level. Intuitively, CUBIC and Westwood protocols might better perform than TCP Newreno/ECN due to the nature of the information returned by the ECN binary signal which does not provide any quantitative estimation of the congestion level allowing TCP to efficiently adapt its sending rate.[1] In other words, whatever the number of ECN marked, the TCP reaction is to halve the congestion window and this action is not well adapted to all cases. Secondly, CUBIC and Westwood are pure end-to-end solutions and as a result, are much more easier to deploy while TCP/ECN must involve both the core network and the end-hosts. However, several research work demonstrate that the design of a mechanism to optimally manage network congestion and capacity while being fair with other flows cannot be done without network collaboration [12,4,6]. Unfortunately and to the best of our knowledge, the major barrier is that we do not have today a solution, that do not involve complex computation inside the core routers (such as BMCC [12] or XCP [6]), able to assess at the sender side the exact congestion level of the bottleneck of the path allowing a transport protocol such as TCP to correctly react to this congestion. For instance, BMCC introduces complex mechanisms inside the router and is only compliant with IPv4 (due to the use of the 16 bits `IPid` field of the IPv4 header) while XCP involves large architectural changes.

This fact motivates the present study which proposes a statistical algorithm to assess the congestion level at the end-hosts side (i.e. receiver or sender sides) without involving complex computation inside the core network. In particular, we aim at providing a practical solution to return concrete congestion measurements to the sender in order to avoid blind, approximate or excessive reaction from the source. The only modification deals with the marking method which is changed from a binary field to a count field similar to the TTL field from the IP packet. Practically, we do not have to extend the IP headers as the DiffServ Codepoint field is large enough to enable our proposal. We could argue, as in [12], whether such modification involves or not heavy IETF standardization process, however we claim that it would be much more complex and uncertain to convince networking companies to add complex estimation method inside their own routers. Furthermore, this solution is generic enough to consider, as for ECN, this flag either as a simple binary indicator or as a counter. Finally, we point out that a recent IETF group named ConEx (Congestion Exposure) [11], attempts to enable congestion to be exposed within the network layer of the Internet. The main candidate solution is to date re-ECN [1] and propose the use of a second bit inside the IP header in order to differentiate the congestion upstream and downstream from an observation point inside the network. Internet service providers are pushing this idea as this would provide an essential tool (currently missing) to better manage and control their traffic.[2] If this solution is adopted, we could assist to a larger deployment of the ECN field that would facilitates the deployment of our proposal.

Following this new marking scheme, we propose a simple method which permits an accurate estimation of the congestion level experienced inside the routers of a given path. We first present the mathematical basis of our proposition then, we provide simulations and the practical analysis to evaluate the congestion level. Finally, we discuss and conclude about the possibility offered by this solutions and detail the remaining work.

2. Marking proposal

The ECN bit, as defined in RFC 3168, is a binary field of the IP header. This field can only contain a boolean value which informs a sender if a packet has crossed at least one congested router. Thus, it is impossible to distinguish a packet marked one time from those marked several times and which would have crossed several congested routers. This prevents any accurate metrology analysis of the link observed for the sake, for instance, of an adapted reaction from the source. In fact, an ECN-capable packet crossing a link composed by two routers and respectively marking at 30% and 40% will have a probability to be marked of 58% (i.e. $1-(1-0.4)(1-0.3)$). Obviously, this does not reflect the level of congestion of the network bottleneck (in this example: 40%) and could lead to an excessive reaction from the source. Thus, we propose to enhance the information returned with an incremental field (denoted ECN^*) to count how many times a packet is marked. The marking scheme, as for RED/ECN, strictly follows the RED algorithm [5]. We will use this new metric (i.e. how many times a packet is marked) to determine the level of congestion of the bottleneck. A RED/ECN^* router will increment this counter instead of simply setting the ECN field to one. Through the analysis of the data received, a source can build the distribution of the marked packets. Obviously, we cannot use this metric as it stands, in the following, we present the analytical method to interpret the data collected.

3. Analytical Study

We present in this part the statistical analysis allowing us to process the data collected with our marking proposal. The results obtained allow to establish a relationship between the frequency of ECN^* marked packets and the queue size of routers of the path.

3.1. Hypotheses and notations

We consider a topology of n core routers in a row. For $1 \leqslant i \leqslant n$, we note R_i the router number i. All these n routers adopt the previously exposed ECN^* marking scheme. Each router drops packets only if its queue is full and probabilistically marks a packet following the average queue size. We call "marking rate" this probability and we adopt the following notations:

[1] We remark that there is a lack of performances evaluation study between ECN-compliant protocols and new proposals such as CUBIC for instance. At least, a recent study clearly shows a clear disequilibrium between TCP Newreno and CUBIC [14].

[2] See the IETF `[re-ecn]` mailing-list and [11] for further details.