Available online at www.sciencedirect.com
SCIENCE DIRECT®
Computer Communications 29 (2006) 525–537
computer communications
www.elsevier.com/locate/comcom

A simulation-based comparative evaluation of transport protocols for SIP

M. Lulling*, J. Vaughan

Department of Computer Science, University College Cork, Western Road, Cork, Ireland

Available online 10 May 2005

Abstract

Research presented in this paper uses the Network Simulator, ns2, to investigate the direct effects and subsequent consequences associated with the use of different transport protocols in a SIP context. Specifically, we seek to perform a comparative evaluation of UDP, TCP and SCTP based on a model simulating a scenario most likely to benefit from congestion control and error correction mechanisms associated with the reliable transport protocols. The overall aim of this paper is to make a contribution to the information available for the appraisal of such protocols in respect of real world implementations.

Keywords: SIP; Transport; TCP; UDP; SCTP; Performance evaluation

Contents

1. Introduction . . . 526
2. Transport for SIP . . . 526
   2.1. SIP over TCP . . . 526
      2.1.1. TCP Reno . . . 527
      2.1.2. TCP Vegas . . . 528
      2.1.3. TCP Sack . . . 528
   2.2. SIP over UDP . . . 529
   2.3. SIP over SCTP . . . 529
3. Simulations . . . 530
   3.1. Induced packet loss . . . 530
   3.2. Random packet loss . . . 530
   3.3. Competing traffic . . . 531
   3.4. Throughput analysis . . . 531
4. Results . . . 531
   4.1. Induced packet loss . . . 531
   4.2. Random packet loss . . . 532
   4.3. Competing traffic . . . 534
   4.4. Throughputs . . . 535
5. Conclusion . . . 536
Acknowledgements . . . 537
References . . . 537

* Corresponding author. Tel.: +353 21 4902795; fax: +353 21 4274390.
E-mail addresses: m.lulling@cs.ucc.ie (M. Lulling), j.vaughan@cs.ucc.ie (J. Vaughan).

0140-3664/$ - see front matter
doi:10.1016/j.comcom.2005.04.006

1. Introduction

This paper presents results from simulation-based experiments to provide information and analysis on the effects of network conditions on Voice over IP (VoIP) signalling traffic and the choice of underlying transport protocol. In this regard the use of SCTP, TCP and UDP in a session initiation protocol (SIP) [1], VoIP signalling context will be compared. As will be detailed in the following sections of this paper, the multiple SIP session scenario present in proxy-to-proxy communication provides an underlying topology that is most likely to benefit from the flow control and congestion control and avoidance features in TCP and SCTP. Results will be presented in terms of delays as well as measured throughput, a performance evaluation gauge first used in this context in preliminary work on the subject [2].

Previous work focussed on the Tahoe, Reno and Sack variants of the TCP transport protocol [3]. It was concluded that Tahoe was completely inappropriate for a SIP signalling context. In this paper, the scope of the research has been extended to include new results for SCTP and UDP as well as the previously omitted Vegas TCP variant. The experimental basis has also been broadened for the simulations using a novel metric to give a further perspective for analysis. Preliminary results comparing TCP Sack to SCTP were presented in [3] and this is comprehensively explored here.

The remainder of this paper is structured as follows. Section 2 discusses transport layer protocols for SIP, specifically TCP, SCTP and UDP. Section 3 introduces the simulations and Section 4 presents the results. Concluding remarks are included in the final section.

2. Transport for SIP

SIP signalling messages are exchanged between IP Telephony users before, during and at the conclusion of a particular IP Telephony session. In this work, the initial messages in session setup are considered. These messages are typically sent from the sender to the receiver via the sender's local SIP proxy and the receiver's local SIP proxy, respectively, as shown in Fig. 1 (provisional responses have been excluded). Subsequent messages are exchanged directly since location information is swapped in the first exchange. Messages exchanged between individual users and their local proxies are characterised as low volume, short-lived, single session messages. Messages from multiple users exchanged between proxies are typically high volume, independent session messages in a long-lived communication session. For this performance evaluation the latter scenario has been focused on, in order to allow the flow control and error correction and recovery mechanisms employed by the TCP and SCTP transport protocols to be used to maximum advantage.

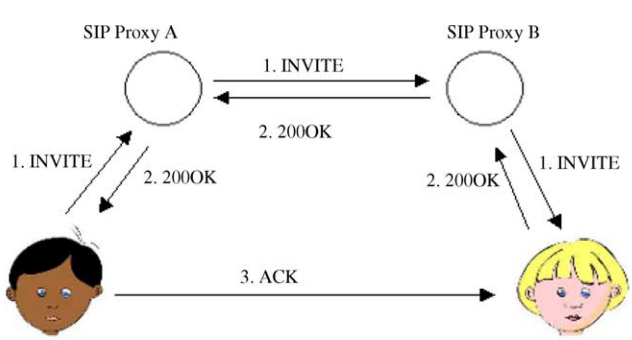

Fig. 1. Typical SIP message sequence for initial call setup.

SIP signalling messages require a delivery guarantee. Due to their nature, SIP messages carry session critical information and thus some mechanism must be in place to ensure that they reach their destination in the event of message corruption or message loss due to network congestion; the latter being the more likely event. UDP does not provide any mechanism in this regard and thus the application layer is responsible for monitoring message status and retransmitting where necessary. Both TCP and SCTP on the other hand take care of this, with the result that the application can 'drop' the message onto the transport layer and 'forget' about it. It would seem, therefore, that a reliable protocol such as TCP or SCTP would be the natural choice; however, both SCTP and TCP have limitations that are particularly pertinent to time sensitive SIP signalling messages. The following subsections will consider SIP transport using the three transport protocols: UDP, TCP and SCTP. For TCP, three variants will be considered, namely the Reno, Vegas and selective Acknowledgement variants.

2.1. SIP over TCP

In the current public Internet, the most widespread transport protocol employed is the transmission control protocol. The reasons for this, historical or otherwise, are outside the scope of this paper, but it is worth noting that a majority portion of competing traffic likely to be encountered on the Internet will be using TCP as a transport protocol. It is also worth noting that TCP is not a concluded standard protocol—there are many different variants with distinguishing differences between most of them, and new versions have been continuously proposed to try to elevate transmission performance.

TCP was not designed with signalling in mind, and has some limitations in this regard. TCP was initially designed to transport large amounts of non-real-time bulk data between two endpoints. A connection is set up between the endpoints and TCP implements flow control and error correction and recovery based on the dynamic behaviour of the end-to-end traffic. Signalling, however, does not typically consist of large amounts of data. SIP messages are usually relatively small (approximately 512 bytes) and SIP uses these messages in a request/response model. Thus, we have small messages that are critically interdependent.