

Chapter 1

A Comparative Study of SIP Overload Control Algorithms

Yang Hong

Carleton University, Canada

Changcheng Huang

Carleton University, Canada

James Yan

Carleton University, Canada

ABSTRACT

Recent collapses of SIP servers in the carrier networks indicates two potential problems of SIP: (1) the current SIP design does not easily scale up to large network sizes, and (2) the built-in SIP overload control mechanism cannot handle overload conditions effectively. In order to help carriers prevent widespread SIP network failure effectively, this chapter presents a systematic investigation of current state-of-the-art overload control algorithms. To achieve this goal, this chapter first reviews two basic mechanisms of SIP, and summarizes numerous experiment results reported in the literatures which demonstrate the impact of overload on SIP networks. After surveying the approaches for modeling the dynamic behaviour of SIP networks experiencing overload, the chapter presents a comparison and assessment of different types of SIP overload control solutions. Finally it outlines some research opportunities for managing SIP overload control.

INTRODUCTION

Internet telephony is experiencing rapidly growing deployment due to its lower-cost telecommunications solutions for both consumer and business services. Session Initiation Protocol (SIP) (Rosenberg et al., 2002) has become the main

signaling protocol to manage multimedia sessions for numerous Internet telephony applications such as Voice-over-Internet Protocol (IP), instant messaging and video conferencing. 3rd Generation Partnership Project (3GPP) has adopted SIP as the basis of its IP Multimedia Subsystem (IMS) architecture (3GPP WG, 2011). With the 3G (3rd Generation) wireless technology being adopted by more and more carriers, most cellular phones and

DOI: 10.4018/978-1-4666-1888-6.ch001

other mobile devices are starting to use or are in the process of supporting SIP for multimedia session establishment (Faccin, Lalwaney, & Patil, 2004).

Figure 1 illustrates a simplified architecture of a SIP network. A SIP network consists of two types of basic elements: User Agent (UA) and Proxy Server (P-Server) (Rosenberg et al., 2002). A user agent can act as a user agent client (UAC) or as a user agent server (UAS). A P-server not only acts as the contact point with UA for core network service access but also provides routing for the signaling messages. SIP is responsible for establishing, modifying and terminating sessions for multimedia communication among multiple UAs.

RFC 5390 (Rosenberg, 2008) identifies the various reasons that may cause server overload in a SIP network. These include but are not limited to poor capacity planning, dependency failures, component failures, avalanche restart, flash crowds, denial of service attacks, etc. In general, anything that may trigger a demand burst or a server slowdown can cause server overload and lead to overload propagation and server crash, thus bringing down the whole SIP network.

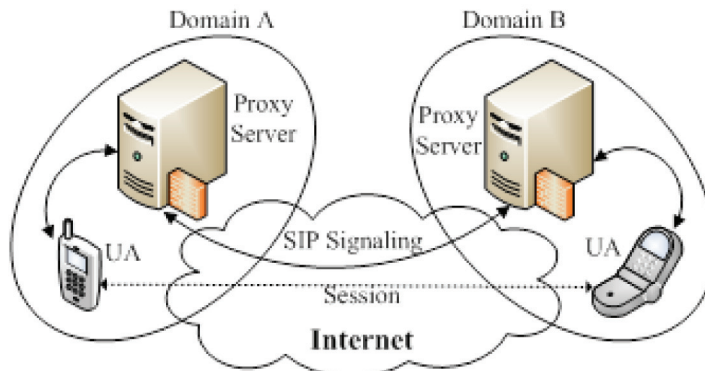
The objective of this chapter is to present a systematic investigation of current state-of-the-art SIP overload control algorithms which aim at preventing server crashes in carrier networks. In order to provide a better knowledge of the major

cause of SIP network collapse, the next section reviews two basic mechanisms of SIP, and describes the existing works on the performance study of the SIP overload. The third section surveys the related SIP modeling and analysis which can help network planners, operators, and researchers to understand how server overloading and widespread SIP network failure may happen under short-term demand bursts or server slowdowns. The fourth section makes a comparative study of different types of SIP overload control solutions, thus helping carriers choose the appropriate solutions to avoid potential SIP network collapse (e.g., Skype outage (Ando, 2010) or VoIP outages in British Telecom, Vonage and Wanadoo (Materna, 2006)) in different overload situations. Finally, some future works for the SIP overload control are discussed.

SIP OVERVIEW

SIP works in the application-layer for session establishment. Figure 2 depicts a typical procedure of a session establishment. To set up a call, a UAC sends an “Invite” request to a UAS via the two proxy servers. The proxy server or the UAS returns a provisional “100Trying” response to confirm the receipt of the “Invite” request. The UAS returns an “180Ring” response after

Figure 1. Simplified architecture of a SIP network



18 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/comparative-study-sip-overload-control/67496

Related Content

Role of Telecommunications in Precision Agriculture

James M. McKinion (2010). *Networking and Telecommunications: Concepts, Methodologies, Tools, and Applications* (pp. 834-848).

www.irma-international.org/chapter/role-telecommunications-precision-agriculture/49779

Introduction of Vehicular Network Applications

Yao-Chung Chang (2010). *Telematics Communication Technologies and Vehicular Networks: Wireless Architectures and Applications* (pp. 15-31).

www.irma-international.org/chapter/introduction-vehicular-network-applications/39517

A Brief Study on Smart Medicine Dispensers

Dayananda P., Amrutha G. Upadhya, Nayana B. G., Priyam Poddarand Vandana Rao Emaneni (2022). *International Journal of Hyperconnectivity and the Internet of Things* (pp. 1-7).

www.irma-international.org/article/a-brief-study-on-smart-medicine-dispensers/294893

Adaptive IoT Technology for Measuring Salinity, Dissolved Oxygen, and pH in Aquatic Environments

Jarrod Trevathanand Dzung Nguyen (2022). *International Journal of Hyperconnectivity and the Internet of Things* (pp. 1-20).

www.irma-international.org/article/adaptive-iot-technology-for-measuring-salinity-dissolved-oxygen-and-ph-in-aquatic-environments/294894

A Survey of IPv6 Addressing Schemes for Internet of Things

Gyanendra Kumarand Parul Tomar (2018). *International Journal of Hyperconnectivity and the Internet of Things* (pp. 43-57).

www.irma-international.org/article/a-survey-of-ipv6-addressing-schemes-for-internet-of-things/221334