

## Research Article

# Prioritizing Signaling Information Transmission in Next Generation Networks

Jasmina Baraković,<sup>1</sup> Himzo Bajrić,<sup>1</sup> Mladen Kos,<sup>2</sup> Sabina Baraković,<sup>3</sup> and Amir Husić<sup>3</sup>

<sup>1</sup> BH Telecom, Joint Stock Company, Sarajevo, Obala Kulina bana 8, 71000 Sarajevo, Bosnia and Herzegovina

<sup>2</sup> Faculty of Electrical Engineering and Computing, University of Zagreb, Unska 3, 10000 Zagreb, Croatia

<sup>3</sup> Department for Informatics and Telecommunication Systems, Ministry of Security of Bosnia and Herzegovina, Trg BiH 1, 71000 Sarajevo, Bosnia and Herzegovina

Correspondence should be addressed to Jasmina Baraković, [jasmina.barakovic@bhtelecom.ba](mailto:jasmina.barakovic@bhtelecom.ba)

Received 1 March 2011; Accepted 1 July 2011

Academic Editor: Akhtar Kalam

Copyright © 2011 Jasmina Baraković et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Next generation transport network is characterized by the use of in-band signaling, where Internet Protocol (IP) packets carrying signaling or media information are mixed in transmission. Since transport resources are limited, when any segment of access or core network is congested, IP packets carrying signaling information may be discarded. As a consequence, it may be impossible to implement reachability and quality of service (QoS). Since present approaches are insufficient to completely address this problem, a novel approach is proposed, which is based on prioritizing signaling information transmission. To proof the concept, a simulation study was performed using Network Simulator version 2 (ns-2) and independently developed Session Initiation Protocol (SIP) module. The obtained results were statistically processed using Statistical Package for the Social Sciences (SPSS) version 15.0. Summarizing our research results, several issues are identified for future work.

## 1. Introduction

Quality of service (QoS) [1] has become one of the most important issues in the next generation network (NGN). Different from the original research on the single node mechanism, current focus of QoS study has shifted to the dynamic end-to-end QoS control [2]. Several standards bodies define the QoS control architectures based on their scope of work such as CableLab, Digital Subscriber Line (DSL) Forum, MultiService Forum (MSF), International Telecommunication Union-Telecommunication Sector (ITU-T), and European Telecommunications Standards Institute (ETSI). ETSI Resource and Admission Control Subsystem (RACS) provides the general architecture covering both access and core networks. There are still many open issues, and continuing effort is under way to solve the issues.

QoS control in the transport technology-dependent aspect is one issue. The standardized NGN QoS concept comes along with a certain volume of signaling information,

which is not efficiently controllable by the respective NGN provider [3]. Because signaling and media information flows of NGN service are mixed while being transmitted, the signaling information flow, which contains a service request signaling packets of the user, may be discarded when any node or any segment of the Internet Protocol (IP) network is overloaded or congested. If the NGN service signaling information flow is unable to be transmitted reliably, it is difficult to guarantee service reachability and implement dynamic end-to-end QoS control. Thus, signaling information transmission is of the most importance in the discussion of the QoS problem in the NGN.

The aim of this paper is to extend our research findings on signaling information transmission in different network conditions. This paper is a continuation of our previous work that introduced the novel approach to configuring Signaling service class [4]. The novel approach is based on using priority queuing system to give the Signaling service class absolute preferential treatment over all other User service classes. While the novel approach is signaling

protocol independent, it is discussed in the context of Session Initiation Protocol (SIP) [5]. Our previous work outlines the issues arising with transmission of SIP signaling flows in case of network congestion [6], bottleneck link overload [7], and link failure [8]. This paper presents an extension of our previous work by attempting to provide the statistical analysis of performance metrics for SIP signaling flows.

The paper is organized as follows. Related work on QoS evolution, signaling, and control is discussed in Section 2. Section 3 summarizes the advantages of the novel approach to signaling information transmission proposed to overcome the existing advantages of standard approach. The simulation study described in Section 4 is performed to verify the novel approach. Section 4.2 presents the obtained results. Section 5 gives concluding remarks and identifies open issues for future work.

## 2. Continuing Evolution of QoS

Providing QoS guarantees in IP networks is a challenging task. First studies proposing QoS frameworks for IP networks started to appear within Internet Engineering Task Force (IETF). To support QoS in IP networks, IETF proposed two frameworks. These are integrated services (IntServ) based on connection-oriented resource reservation principle and differentiated services (DiffServ) based on service differentiation approach. IntServ provides deterministic QoS guarantees and requires a signaling protocol in order to inform network elements about the required reservation. Resource Reservation Protocol (RSVP) is a protocol proposed for IntServ. The limitations of RSVP have inspired the emerging Next Steps in Signaling (NSIS) framework being developed by the IETF, for the purpose of installing and maintaining flow states in the network. The building blocks relevant to the IntServ framework include admission control, queuing, resource reservation, traffic classification, and traffic policing.

The concept behind the DiffServ framework is treating a packet based on its class of service as encoded in its IP header [9]. DiffServ defines the structure and meaning of the type of service (ToS) byte in the IP header. Within the DiffServ framework, this byte is called differentiated services (DS) field [10]. Its value is called DiffServ code point (DSCP). Based on this field, DiffServ specification defines a number of packet forwarding treatments to be implemented in network elements. They are called per-hop Behaviors (PHB). In contrast to the IntServ framework, the treatment is based not on a per-flow basis, but solely on the indicated class of service. The building blocks relevant to the DiffServ framework include buffer management, packet marking, service level agreement (SLA), traffic metering and recording, traffic policing, traffic shaping, and scheduling.

Although multiprotocol label switching (MPLS) is not considered as a QoS framework for IP networks, it provides a number of advantages in terms of QoS support. Modern QoS-aware networks are based on the combination of DiffServ and MPLS [11]. Such networks are designed to be flexible enough to reallocate network resources in the best possible way, such that the required performance is

provided using minimum amount of resources. Whereas DiffServ/MPLS network is strong in delivering scalable multiclass services, it is still weak in providing standardized methods to automatically describe arriving traffic and dynamically reallocate resources. These issues are covered under the scope of QoS control, which can be implemented in many different ways.

The QoS control architectures have been developed in several standards bodies [2], in which the study of the Telecommunications and Internet Converged Services and Protocols for Advanced Networking (TISPAN) workgroup of ETSI wins the recognition of the industry. Due to diversity and multimedia feature of NGN services, how to implement QoS control on user service on the basis of IP-based transport layer has become an important task in NGN technical research.

In the NGN, a critical component that provides support for the dynamic QoS control is RACS [12]. The RACS needs to have interfaces with transport functional (TF) entity, service control function (SCF) entity, and network attachment subsystem (NASS) in order to enable dynamic end-to-end QoS control depending on the availability of the current network resources and thereby to improve network resource utility effectively.

Based on the RACS and the interface relationship between the RACS and the external entities in the NGN, it is possible to support different QoS control scenarios for the services requested by user terminals with different QoS negotiation capabilities [13]. Three different NGN QoS control scenarios can be distinguished according to ETSI/TISPAN [3]. The application of a specific scenario depends on the QoS signaling capabilities of the respective user equipment that initializes a session requiring certain QoS conditions within the respective NGN transport network.

User equipment can be classified into the following categories: (1) without QoS negotiation capabilities, (2) with QoS negotiation capabilities at transport layer such as RSVP or NSIS, (3) and with QoS negotiation capabilities at service layer such as SIP. In the current network environments, most of the user terminals do not support RSVP or NSIS protocol, because SIP has been adopted by the telecommunications industry as its protocol of choice for signaling. Thereby, during the service setup process, the user terminal can negotiate the QoS parameters required for the current service with the opposite terminal or application server via SIP. After the user terminal obtains the QoS parameters through negotiation, it sends a service request carrying the QoS parameters to the SCF entity. When the SCF entity receives the user service request, it extracts the QoS parameters from the service request and then forwards the QoS request to the RACS. Based on user profile, operator-specific policy rules, and resource availability, the RACS performs authentication and makes admission control decision on the received QoS request. When the RACS determines to permit the service to be transferred in the network in accordance with the requested QoS parameters, it sends gate control, bandwidth allocation, aggregation and adaptation control commands, and so forth, to the TF entity at the boundary of the network for traffic forwarding.

### 3. Signaling Information Transmission for QoS Control

Because of the fact that both service and transport layers are involved in the NGN QoS control process, the signaling between these two layers is compulsory. Combined with the application of traditional IP QoS mechanisms, this leads to a certain volume of signaling traffic [3]. The signaling traffic has always been considered as negligible and treated as best-effort traffic. This occurs even in many QoS-related projects, except End-to-end Quality of Service support over heterogeneous networks (EuQoS) project [14]. If signaling traffic is not treated properly while being transmitted over IP network, NGN QoS control may not work efficiently. Moreover, service reachability may fail to be implemented. For implementing service reachability, signaling information flow of NGN services should be ensured to be transmitted instantly and reliably.

*3.1. Standard Approach to Signaling Information Transmission.* In an all-IP transport network, a common interface between a user and a network utilizes the IP technology. In an IP access network of the NGN, all IP packets issued by a user are transmitted forward with a best-effort mode or a priority mode and signaling and media information flows of the NGN services are mixed in transmission. Since network resources are always limited, when any segment or any node of the network between a user and a SCF entity is overloaded or congested, the signaling information flow of the NGN services may be discarded and unable to reach the SCF entity. As a result, even if the user requests services of a relatively higher priority, it is possible that no answer will be received and the service reachability as the first ingredient in the QoS perceived by the user fails to be implemented. Moreover, in the case of mixing the signaling and media information flows in transmission, it is difficult for an IP network transmission node to identify whether the transmitted IP packet contains a signaling information or a media information and to guarantee that the IP packets containing a signaling information are not discarded in the case of the congestion, which makes it obviously impossible to implement the service reachability.

The service requested by a user may also be an emergent call service with a top priority such as a burglar alarm call, a fire alarm call, a traffic alarm call, or a first aid call. Such a service requires not only that the signaling information flow of service can always be transmitted instantly and reliably, but also that the media information flow of service can be established on demand and be transmitted reliably. If a signaling information flow is unable to reach a SCF entity reliably, it is impossible to establish an application session and start a media information flow.

Under normal network conditions, national and international networks can meet the need of NGN services, except in the less-developed regions of the world where bandwidth is still badly lacking [15]. The network performance statistics suggest that national networks in developed countries should already be providing reachability and quality of NGN services. However, it is generally felt that network performance

can be unpredictable at times. There is a question how to explain this contradiction in network performance prediction. Most NGN providers have chosen to overprovision their network bandwidth based on the statistics that is, obtained from traffic monitoring systems. However, these systems monitor traffic with large time intervals. Due to averaging over large time intervals, network links seem to be underutilized, when, in fact, at time granularity smaller than a second, performance degradation is present leading to unacceptable QoS. The link utilization at various time scales (millisecond level) is analyzed to measure the frequency and duration of microcongestion, which is caused by bandwidth reduction, link multiplexing, and traffic burstiness. Analysis reveals that packet losses on underutilized link do occur due to microcongestion events [16].

To sum up, it is not possible to guarantee that signaling information flow of the NGN services is transmitted instantly and reliably, which makes it difficult to implement the service reachability. The signaling information flow of the NGN services is unable to reach the SCF entity, so that the media information flow cannot be transmitted, which makes it difficult to implement an emergent call service and trigger the RACS.

*3.2. Novel Approach to Signaling Information Transmission.* The NGN provider should guarantee that the signaling information flow of all NGN services is transmitted instantly and reliably regardless of the service level or the service type. It has not been recognized that guaranteeing the service reachability, that is, the instant and reliable transmission of signaling information flows of all telecom-level services in any case of resource occupancy rate, is of the most importance in the discussion of the QoS problem of the NGN.

To accomplish the task of prioritizing signaling information transmission, DiffServ addresses the clear need for relatively simple and coarse methods of categorizing traffic into various service classes and applying QoS parameters to those classes. The IETF has released Request for Comments (RFC) 4594, which proposes configuration guidelines for DiffServ service classes [17]. Since these guidelines are not standards, modifications can be made to these recommendations as required by specific needs and constraints.

Different service classes are constructed using DSCPs, traffic conditioners, PHBs, and active queue management (AQM) mechanisms. The Signaling service class should be configured to provide a minimum bandwidth assurance for Class Selector 5 (CS5) marked packets as defined in RFC 4594. This service class should use a rate queuing system, such as weighted fair queuing (WFQ) or weighted round robin (WRR). The single rate with burst size token bucket policer should be used to ensure that signaling traffic stays within its negotiated or engineered bounds. Since the traffic in this service class does not respond dynamically to packet loss, AQM should not be applied to CS5 marked packets [17]. To meet specific business requirements, Cisco has made a minor modification to its adoption of RFC 4594, namely, the switching of Signaling and Broadcast Video service class markings (to CS3 and CS5, resp.) [18].

As defined in RFC 4594, the signaling service class should be configured to assure the target values for IP packet transfer delay (IPTD) and IP packet loss ratio (IPLR), while the value of IP packet delay variation (IPDV) is not crucial. According to ITU-T Recommendation Y.1541, the following target values should be guaranteed: upper bound on IPTD below 100 ms and upper bound on IPLR below  $1 \times 10^{-3}$ . To meet these QoS requirements regardless of network conditions, we proposed a modification of the configuration guideline for Signaling service class [4]. The modification consists of Signaling service class configuring by using priority queuing system to give it absolute preferential over all other User service classes. The priority queuing system is a combination of a set of queues and a scheduler that empties them in priority sequence [17]. When asked for a packet, the scheduler inspects the highest priority queue dedicated to Signaling service class and, if there is traffic present, returns a packet from that queue. Failing that, it inspects the next highest priority queue, and so on. A packet assigned to Signaling service class should be marked with a new DSCP value, which should be requested from Internet Assigned Numbers Authority (IANA). This DSCP should be lower than one used to configure the Network Control service class and higher than one reserved for all User service classes defined in RFC 4594 and RFC 5865 [19].

Applying the proposed modification to Signaling service class configuration, the QoS objectives for signaling traffic could be guaranteed regardless of the network conditions. This possibility is essential for efficiency of standardized NGN approaches for QoS control. If signaling traffic is able to reach the SCF entity, it is possible to implement NGN service reachability and trigger RACS by means of service request. As a result, it is possible to enable dynamic end-to-end QoS control depending on the availability of the current network resources and thereby to improve transport network resource utility effectively.

## 4. Simulation Investigation

To verify the novel approach to signaling information transmission, the simulation study is performed using a Network Simulator version 2 (ns-2) [20] and independently developed SIP module [21].

*4.1. Simulation Methodology.* Three experiments, which differ from each other according to the cause and localization of congestion, are conducted as shown in Table 1. We refer to Experiment 1 when congestion occurs due to network overload, Experiment 2 when congestion occurs due to bottleneck links overload, and Experiment 3 when congestion occurs due to multiple links failures.

Each experiment includes four different simulation scenarios, which differs from each other according to the forward priority assigned to different traffic flows as shown in Table 1. We will refer to Scenario 1 when SIP signaling flows are transmitted with a best-effort mode and Scenarios 2, 3, and 4 when SIP signaling flows are transmitted with a DiffServ mode. In Scenario 2, the SIP signaling flows share

TABLE 1: Description of experiments and scenarios.

| (a) Experiments                          |                |                       |                       |                       |
|--|----------------|-----------------------|-----------------------|-----------------------|
|  | Experiment 1   | Experiment 2          | Experiment 3          |                       |
| Cause of congestion                      | Overload       | Overload              | Overload Failure      |                       |
| Localization of congestion               | Network        | B → C&D → F           | Network A → B&D → B   |                       |
| (b) Scenarios                            |                |                       |                       |                       |
|  | Scenario 1     | Scenario 2            | Scenario 3            | Scenario 4            |
| QoS <sup>a</sup> Model                   | Best effort    | DiffServ <sup>b</sup> | DiffServ <sup>b</sup> | DiffServ <sup>b</sup> |
| Service class                            | Priority level |                       |                       |                       |
| Telephony (EXP <sup>c</sup> )            |                | 1                     | 1                     | 2                     |
| Multimedia Streaming (CBR <sup>d</sup> ) | No priority    | 2                     | 2                     | 3                     |
| Standard (FTP <sup>e</sup> )             |                | 3                     | 4                     | 4                     |
| Signaling (SIP <sup>f</sup> )            |                |                       | 3                     | 1                     |

<sup>a</sup>QoS: quality of service; <sup>b</sup>DiffServ: differentiated service; <sup>c</sup>EXP: exponential; <sup>d</sup>CBR: constant bit rate; <sup>e</sup>FTP: file transfer protocol; <sup>f</sup>SIP: session initiation protocol.

TABLE 2: Information about traffic paths.

| Path followed                          | $i$ | Number of flows                                |  |  |  |
|--|-----|--|--|--|--|
|  |     | EXPpath <sub><math>i</math></sub> <sup>a</sup> | CBRpath <sub><math>i</math></sub> <sup>b</sup> | FTPpath <sub><math>i</math></sub> <sup>c</sup> | SIPpath <sub><math>i</math></sub> <sup>d</sup> |
| D <sub>1</sub> -A-B-C-D <sub>5</sub>   | 1   | 10   | 16   | 200  | 10   |
| D <sub>1</sub> -A-D-F-D <sub>4</sub>   | 2   | 10   | 16   | 200  | 10   |
| D <sub>2</sub> -B-C-D <sub>5</sub>     | 3   | 10   | 16   | 200  | 10   |
| D <sub>3</sub> -E-D-B-C-D <sub>5</sub> | 4   | 10   | 16   | 200  | 10   |
| D <sub>3</sub> -E-D-F-D <sub>4</sub>   | 5   | 10   | 16   | 200  | 10   |
| D <sub>4</sub> -F-B-C-D <sub>5</sub>   | 6   | 10   | 16   | 200  | 10   |

<sup>a</sup>EXP: exponential; <sup>b</sup>CBR: constant bit rate; <sup>c</sup>FTP: file transfer protocol; <sup>d</sup>SIP: session initiation protocol.

the lowest priority queue with FTP data flows. In Scenario 3, the SIP signaling flow packets are assigned to the dedicated low-priority queue, whose limit on maximum bandwidth is 10% of available link capacity. The highest priority queue is dedicated to the SIP signaling flows in Scenario 4.

The simulations are based on a common network topology consisting of six routers named A, B, C, D, E, and F, with five user domains (D<sub>1</sub>, D<sub>2</sub>, D<sub>3</sub>, D<sub>4</sub>, and D<sub>5</sub>) directly attached to each of them. The network topology is shown in Figure 1. The link capacities are dimensioned to implement a network configuration whose load equals 20%. We first consider a configuration in which each buffer in the network router has a service rate five times greater than the sum of the goodputs required by the traffic flows passing through it. The configured link capacities as well as propagation delays are shown in Figure 1. Starting from this configuration, we decrease the capacity of appropriate links according the total

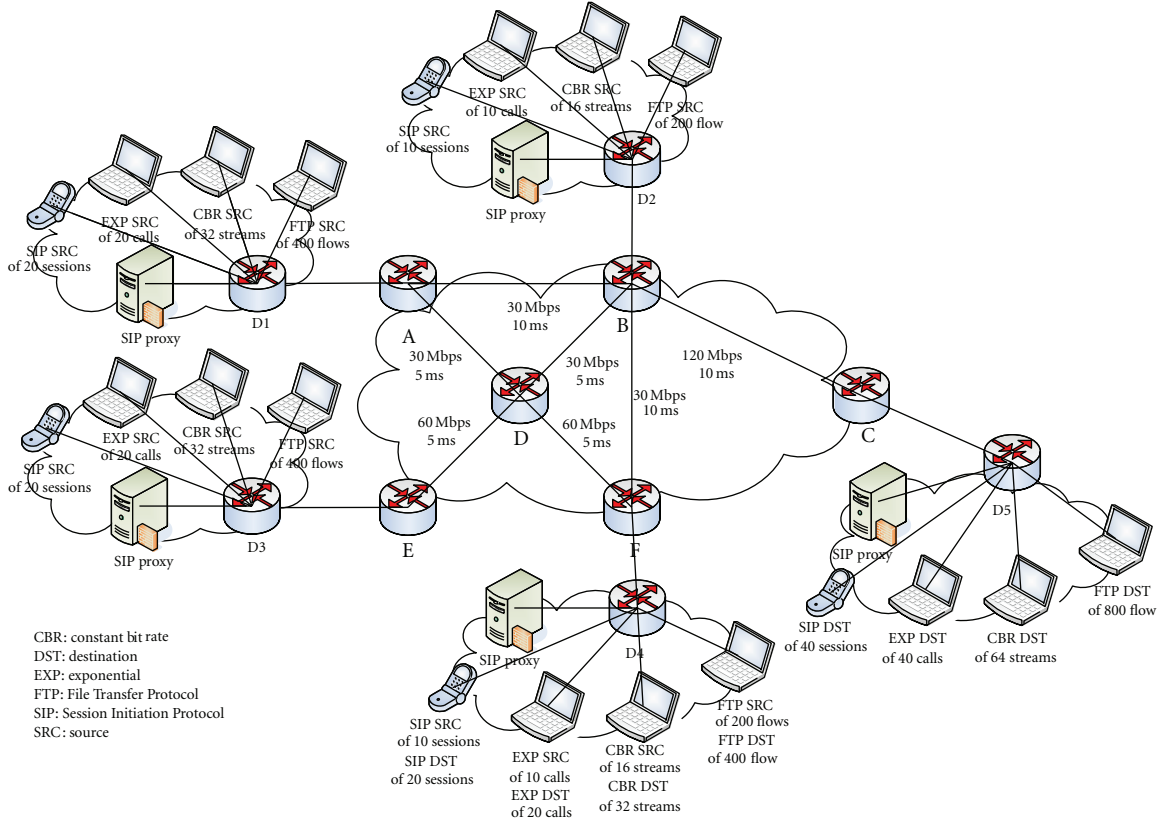


FIGURE 1: Simulation network topology.

TABLE 3: Network/link load percentage above which performance metrics for SIP signaling flows exceed target QoS values.

| Experiment/scenario | Scenario 1        |                   | Scenario 2        |                   | Scenario 3        |                   | Scenario 4        |                   |
|---------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
|                     | IPTD <sup>a</sup> | IPLR <sup>b</sup> | IPTD <sup>a</sup> | IPLR <sup>b</sup> | IPTD <sup>a</sup> | IPLR <sup>b</sup> | IPTD <sup>a</sup> | IPLR <sup>b</sup> |
| Experiment 1        | 60                | 40                | 60                | 60                | 90                | 100               | —                 | —                 |
| Experiment 2        | 80                | 80                | 60                | 100               | 90                | 100               | —                 | —                 |
| Experiment 3        | 20                | 20                | 20                | 20                | 20                | 100               | —                 | —                 |

<sup>a</sup>IPTD: IP transfer delay; <sup>b</sup>IPLR: IP loss ratio.

TABLE 4: Regression model for analyzing IPTD for SIP signaling flows.

| IPTD <sup>a</sup> | Regression model | Parameters          |        | $r^{2b}$              | $r^c$  | $p^d$  |       |
|-------------------|------------------|---------------------|--------|-----------------------|--------|--------|-------|
|                   |                  | $b_0$               | $b_1$  |                       |        |        |       |
| Experiment 1      | Scenario 1       | $Y = b_0 e^{b_1 t}$ | 0.0458 | 0.0199                | 0.0932 | 0.3054 | <0.05 |
|                   | Scenario 2       | $Y = b_0 e^{b_1 t}$ | 0.0385 | 0.0251                | 0.1812 | 0.4257 | <0.05 |
|                   | Scenario 3       | $Y = b_0 e^{b_1 t}$ | 0.0123 | 0.6913                | 0.0384 | 0.1960 | <0.05 |
|                   | Scenario 4       | $Y = b_0 + b_1 t$   | 0.0716 | $1.52 \times 10^{-5}$ | 0.0027 | 0.0516 | >0.05 |
| Experiment 2      | Scenario 1       | $Y = b_0 e^{b_1 t}$ | 0.0243 | 0.0135                | 0.3788 | 0.6155 | <0.05 |
|                   | Scenario 2       | $Y = b_0 e^{b_1 t}$ | 0.0213 | 0.0280                | 0.4314 | 0.6568 | <0.05 |
|                   | Scenario 3       | $Y = b_0 e^{b_1 t}$ | 0.0240 | 0.0167                | 0.3119 | 0.5584 | <0.05 |
|                   | Scenario 4       | $Y = b_0 + b_1 t$   | 0.0613 | $2.34 \times 10^{-6}$ | 0.0027 | 0.0516 | >0.05 |
| Experiment 3      | Scenario 1       | $Y = b_0 e^{b_1 t}$ | 0.0458 | 0.0199                | 0.0932 | 0.3054 | <0.05 |
|                   | Scenario 2       | $Y = b_0 e^{b_1 t}$ | 0.0385 | 0.0251                | 0.1812 | 0.4257 | <0.05 |
|                   | Scenario 3       | $Y = b_0 e^{b_1 t}$ | 0.0938 | 0.0121                | 0.0351 | 0.1873 | <0.05 |
|                   | Scenario 4       | $Y = b_0 + b_1 t$   | 0.0715 | $1.52 \times 10^{-5}$ | 0.0027 | 0.0516 | >0.05 |

<sup>a</sup>IPTD: IP packet transport delay; <sup>b</sup> $r^2$ : determination coefficient; <sup>c</sup> $r$ : correlation coefficient; <sup>d</sup> $p$ : significance.

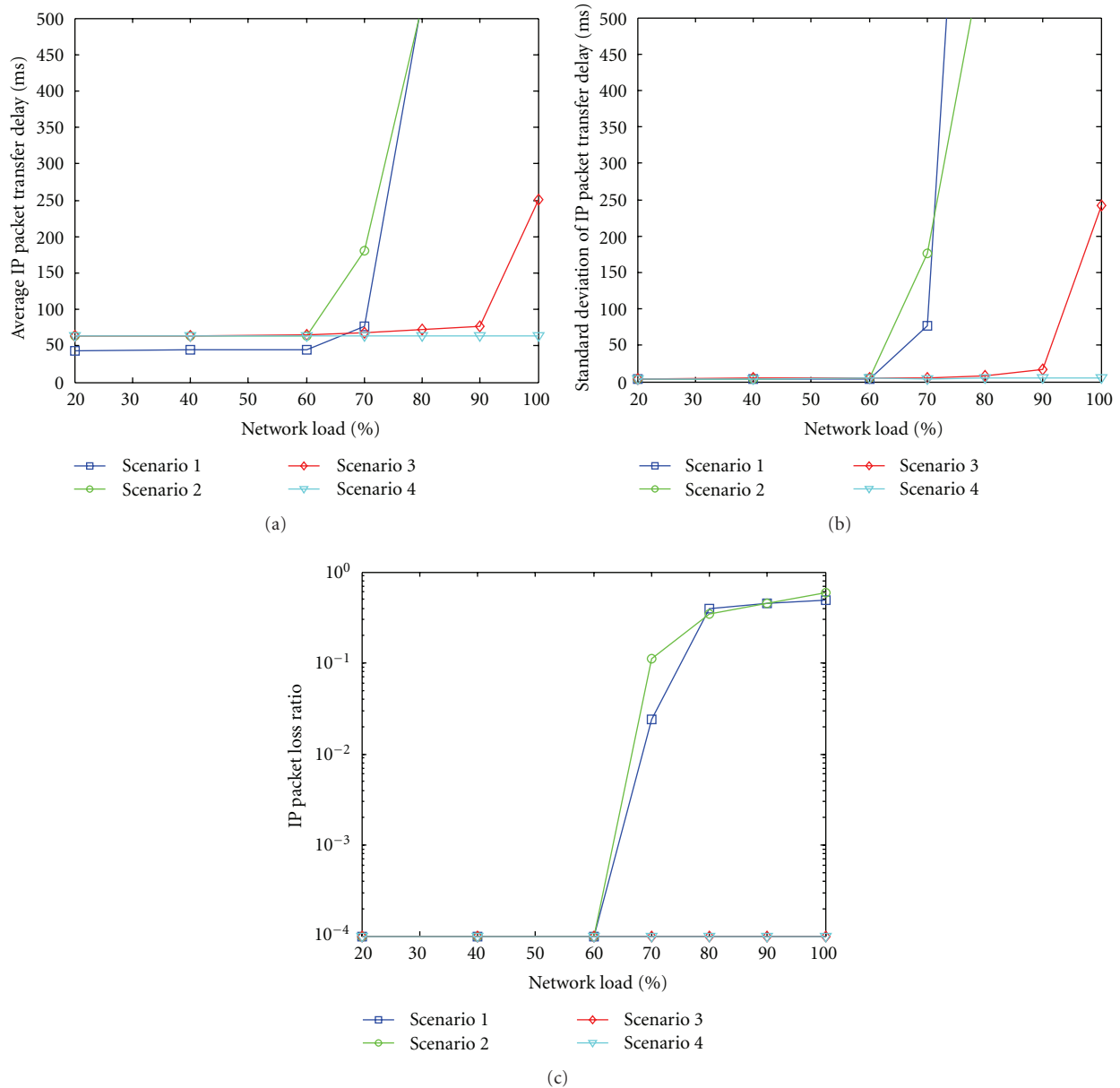


FIGURE 2: Critical performance metrics for SIP signaling flow 1—Experiment 1: (a) average IPTD for SIP signaling flow 1; (b) standard deviation of IPTD for SIP signaling flow 1; (c) IPLR for SIP signaling flow 1.

amount of incoming traffic passing through it, which is kept constant and equals 6 Mb.

The network is loaded by four types of traffic flows, which follow the same paths  $i$  ( $i = 1, 2, 3, 4, 5, 6$ ) listed in Table 2. The Exponential (EXP), Constant Bit Rate (CBR), and File Transfer Protocol (FTP) traffic generators are used to simulate background traffic. The EXP and the CBR traffic generators are attached to User Datagram Protocol (UDP) agents, while the FTP traffic generators are attached to Transport Control Protocol (TCP) agents. EXP traffic generator is configured to generate 200 byte voice packets with burst time of 1 second, idle time of 0.5 second, and generation rate of 100 kbps. The CBR agent is configured to generate 1000 byte video packets and generation rate

300 kbps. FTP traffic generators simulate flows generated by data-limited TCP sources. We use Reno version of TCP sending agent with a maximum window size of 2000- and 1000-byte fixed size packets. The average time between the arrivals of new TCP sessions is 4 seconds. We generate sessions with random size with a mean of  $1 \times 10^5$  byte, with exponential distribution and shape 1.5.

The global generation rate of FTP flows is thus 200 kbps, which equals 3.33% of total amount of generated traffic. By multiplying the number of generated EXP and CBR flows shown in Table 2 and the generation rate of each ones, the global generation rate of EXP and CBR flows is calculated. The global generation rate of EXP and CBR flows equals 1 Mbps and 4.80 Mbps, constituting 16.67% and 80%

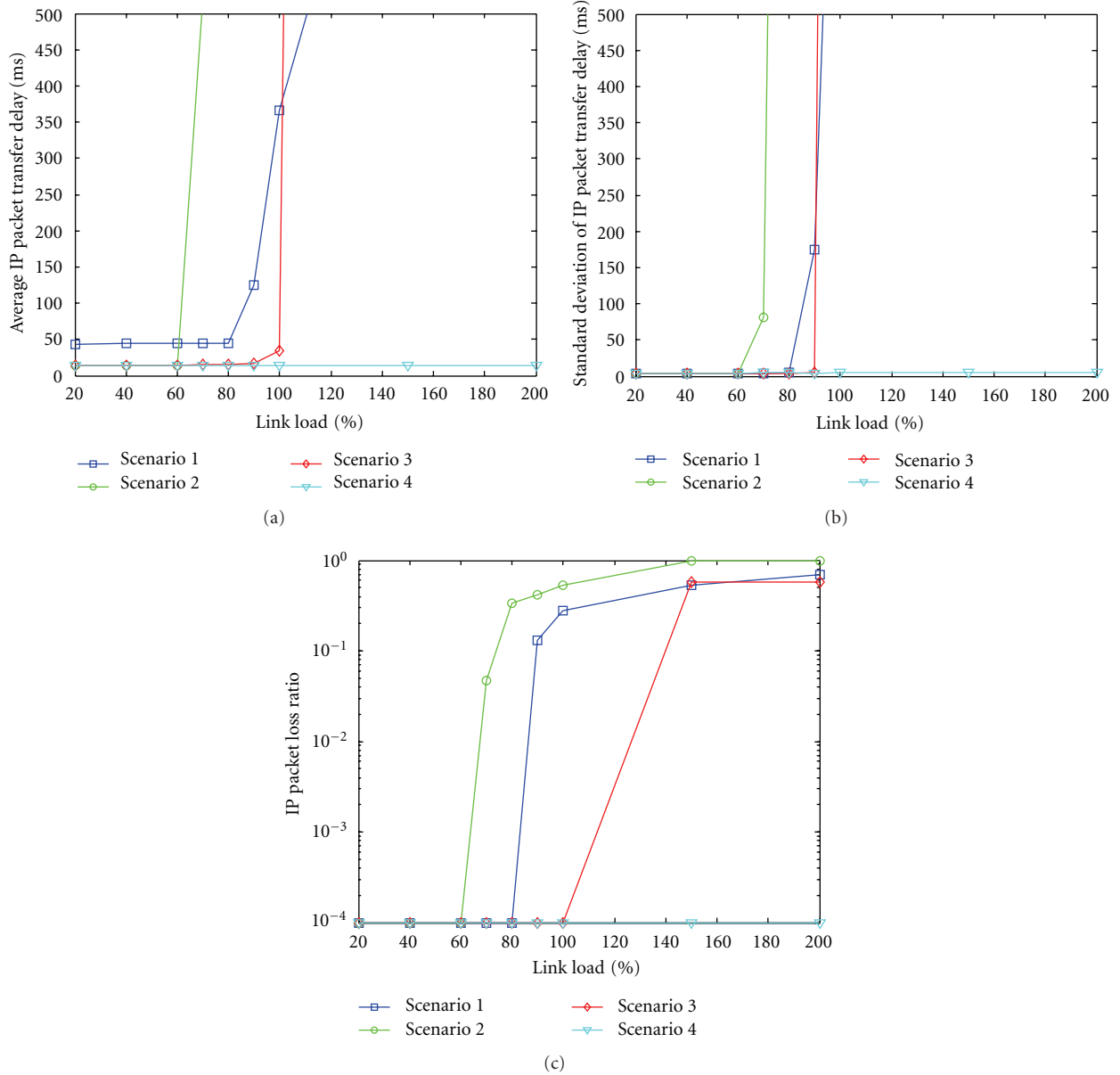


FIGURE 3: Critical performance metrics for SIP signaling flow 1—Experiment 2: (a) average IPTD for SIP signaling flow 1; (b) standard deviation of IPTD for SIP signaling flow 1; (c) IPLR for SIP signaling flow 1.

of total amount of traffic, respectively. Based on practical experiences, SIP signaling flows constitute less than 1% of total amount of traffic generated in the network, and as such cannot cause the congestion.

Scenario 1 is characterized by the absence of QoS mechanisms, while the specific combination of them has been considered in Scenarios 2, 3, and 4. It is assumed that in Scenarios 2, 3, and 4 all the routers adopt the same weighted random early detection (WRED) mechanism and priority queuing (PQ) algorithm. The configuration parameters for each WRED router equal to ns-2 default ones [20]. The PQ algorithm ensures that important traffic gets the fastest handling at each point where it is used. It was designed to give strict priority

to important traffic. In the PQ, each packet is placed in one of four queues—high, medium, normal, or low—based on an assigned priority as shown in Table 1. During transmission, the algorithm gives higher-priority queues absolute preferential treatment over low-priority queues. The Token Bucket policy model is chosen with the aim of correct marking of EXP, CBR, and FTP traffic flows, whereas the Null policy model is applied to SIP signaling flows. The configured Committed Information Rate (CIR) value equals to 2Mbps for both EXP and FTP traffic flows, and 6Mbps for CBR traffic flows. The committed burst size (CBS) value equals to 0.15Mbps for both EXP and FTP traffic flows and 0.45Mbps for CBR traffic flows.

TABLE 5: Regression model for analyzing IPLR for SIP signaling flows.

| IPLR <sup>a</sup> | Regression model | Parameters                         |         |         |         | $r^{2b}$               | $r^c$  | $p^d$  |       |
|-------------------|------------------|------------------------------------|---------|---------|---------|------------------------|--------|--------|-------|
|                   |                  | $b_0$                              | $b_1$   | $b_2$   | $b_3$   |                        |        |        |       |
| Experiment 1      | Scenario 1       | $Y = b_0 + b_1t + b_2t^2 + b_3t^3$ | -0.3573 | -0.0251 | -0.0004 | $-1.73 \times 10^{-6}$ | 0.6858 | 0.8281 | <0.05 |
|                   | Scenario 2       | $Y = b_0 + b_1t + b_2t^2 + b_3t^3$ | 0.1538  | -0.0094 | 0.0001  | $-2.25 \times 10^{-7}$ | 0.8369 | 0.9148 | <0.05 |
|                   | Scenario 3       | —                                  | —       | —       | —       | —                      | —      | —      | —     |
|                   | Scenario 4       | —                                  | —       | —       | —       | —                      | —      | —      | —     |
| Experiment 2      | Scenario 1       | $Y = b_0 + b_1t + b_2t^2 + b_3t^3$ | 0.1734  | -0.0110 | 0.0002  | $-4.58 \times 10^{-7}$ | 0.9485 | 0.9739 | <0.05 |
|                   | Scenario 2       | $Y = b_0 + b_1t + b_2t^2 + b_3t^3$ | 0.1807  | -0.0144 | 0.0003  | $-8.38 \times 10^{-7}$ | 0.9673 | 0.9836 | <0.05 |
|                   | Scenario 3       | $Y = b_0 + b_1t + b_2t^2 + b_3t^3$ | 0.3318  | -0.0176 | 0.0002  | $-6.01 \times 10^{-7}$ | 0.8643 | 0.9296 | <0.05 |
|                   | Scenario 4       | —                                  | —       | —       | —       | —                      | —      | —      | —     |
| Experiment 3      | Scenario 1       | $Y = b_0 + b_1t + b_2t^2 + b_3t^3$ | -0.5489 | 0.0370  | -0.0005 | $2.71 \times 10^{-6}$  | 0.8847 | 0.9405 | <0.05 |
|                   | Scenario 2       | $Y = b_0 + b_1t + b_2t^2 + b_3t^3$ | -1.2363 | 0.0881  | -0.0015 | $-8.04 \times 10^{-6}$ | 0.9682 | 0.9839 | <0.05 |
|                   | Scenario 3       | —                                  | —       | —       | —       | —                      | —      | —      | —     |
|                   | Scenario 4       | —                                  | —       | —       | —       | —                      | —      | —      | —     |

<sup>a</sup>IPLR: IP packet loss ratio; <sup>b</sup> $r^2$ : determination coefficient; <sup>c</sup> $r$ : correlation coefficient; <sup>d</sup> $p$ : significance.

**4.2. Simulation Results and Discussion.** The output trace file from each simulation scenario in Experiments 1, 2, and 3 is used to measure average IPTD and IPLR obtained for SIP signaling flows. To study the distribution of these packets around the mean, the standard deviation ( $\sigma$ ) in the IPTD is calculated. Using standard deviation, it is possible to estimate region around the mean where 68%, 95%, and 99% of IPTDs are located if the distribution is symmetric about the mean.

Experiment 1 investigates effects of network congestion on the critical performance metrics for six SIP signaling flows through four simulation scenarios [6]. The behavior of performance metrics for six SIP signaling flows in case of the bottleneck link congestion is analyzed in Experiment 2 [7]. The effect of a double link failure on critical performance metrics for six SIP signaling flows is studied in Experiment 3 [8]. Although the critical performance metrics are measured for six SIP signaling flows, this paper shows only results obtained for SIP signaling flow 1 (Figures 2, 3, and 4). The network/link load percentage, above which the critical performance metrics for six SIP signaling flows exceed the target values, is determined in each simulation scenario. The results obtained in each experiment are shown in Table 3.

In Experiment 1, the average IPTD exceeds the target value when the network load is above 60% in Scenarios 1 and 2 and above 90% in Scenario 3. The average IPLR exceeds the target value when the bottleneck links load is above 40% in Scenario 1, 60% in Scenario 2, and 100% in Scenario 3. The average IPLR does not exceed  $1 \times 10^{-3}$  regardless of the network load and equals zero only in Scenario 4.

In Experiment 2, the average IPTD exceeds 100 ms when the bottleneck links load is above 80% in Scenario 1, 60% in Scenario 2, and 90% in Scenario 3. The average IPLR exceeds the target value when the bottleneck links load is above 80% in Scenario 1 and 100% in Scenarios 2 and 3. The average IPLR does not exceed  $1 \times 10^{-3}$  regardless of the bottleneck links load and equals zero in Scenario 4.

In Experiment 3, the average IPTD does not exceed 100 ms only in Scenario 4. The average IPLR exceeds the target value when the network load is above 20% both in

Scenarios 1 and 2 and 100% in Scenario 3. It is kept below  $1 \times 10^{-3}$  regardless of the network load and equals zero only in Scenario 4.

Summarizing simulation results, it is noticed that the critical performance metrics for SIP signaling flows is possible to keep under target values regardless of network/link load only when the novel approach to configuring Signaling service class is used (Scenario 4). The average value of IPTD obtained for SIP signaling flows in Scenario 4 equals 60.43 ms ( $\sigma = \pm 6.59$  ms) in Experiment 1, 59.99 ms ( $\sigma = \pm 6.50$  ms) in Experiment 2, and 70.91 ms ( $\sigma = \pm 8.27$  ms) in Experiment 3. The IPLR of 0 obtained for SIP signaling flows is detected in Scenario 4 of each experiment.

The obtained results are statistically processed using Statistical Package for the Social Sciences (SPSS) version 15.0. The null hypothesis is setup stating that the target QoS objectives for Signaling service class are guaranteed when it is assigned the highest priority. Against the null hypothesis is setup the alternative hypothesis. Statistical tests of the null hypothesis are summarized as a  $P$  value with an indication of its statistical significance. The 95% confidence interval is chosen, and this relates to the generally accepted level of statistical significance  $P < 0.05$ . The regression analysis is conducted to find the relationship that explains how the variation in IPTD/IPLR values for SIP signaling flows depends on the variation in a network/link load. While the coefficient of correlation ( $r$ ) is measured to give the true direction of the correlation between the IPTD/IPLR values for SIP signaling flows and network/link load, the coefficient of determination ( $r^2$ ) is provided to give the true strength of this correlation but without an indication of its direction.

The results of regression analysis of the relationship between the IPTD for SIP signaling flows and the network/link load for each scenario of conducted experiments are summarized in Table 4. The exponential regression model is used to evaluate the relationship between IPTD for SIP signaling flows as dependent variable and network/link load as independent variable in the first three scenarios. The linear regression model is used to describe how the typical value



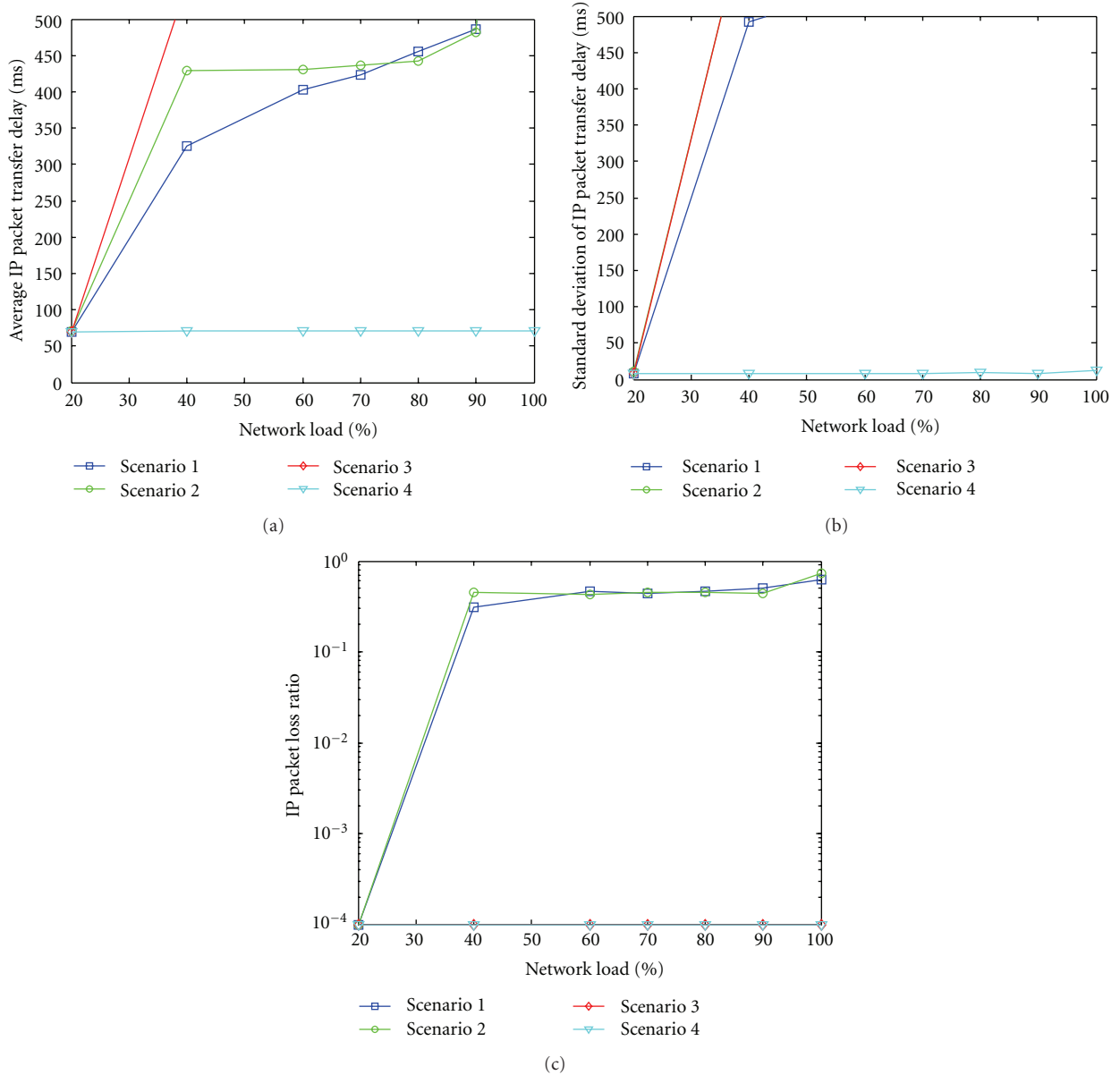


FIGURE 4: Critical performance metrics for SIP signaling flow 1—Experiment 3: (a) average IPTD for SIP signaling flow 1; (b) standard deviation of IPTD for SIP signaling flow 1; (c) IPLR for SIP signaling flow 1.

of IPTD for SIP signaling flows changes when network/link load is varied in Scenario 4. As shown in Table 4, the medium positive correlation between IPTD for SIP signaling flows and network/link load found in Scenarios 1 and 2 of each experiment, as well as in Scenario 3 of Experiment 2, is statistically significant ( $P < 0.05$ ). This is in contrast with results obtained in Scenario 4 of each experiment. The correlation coefficient found in Scenario 4 between the IPTD and the network/link load equals 0.0516 and is not statistically significant ( $P > 0.05$ ). The coefficient of determination equals 0.0027, which means that only 0.27% of the total variation in the IPTD for SIP signaling flows can be explained by the linear relationship between the IPTD and the network/link load. It means that there is almost no statistical relationship between the IPTD and the

network/link load when SIP signaling flows are marked with the highest priority. If the network load increases by 1%, we predict that the IPTD for SIP signaling flows will increase by approximately 0.0152 ms ( $b_1 = 1.52 \times 10^{-5}$ ) in Experiments 1 and 3 and 0.00234 ms ( $b_1 = 2.34 \times 10^{-3}$ ) in Experiment 2. If the network load is 0%, then we predict that the IPTD for SIP signaling flows is 71 ms ( $b_0 = 0.0716$ ) in Experiment 1, 61.3 ms ( $b_0 = 0.0613$ ) in Experiment 2, and 71.5 ms ( $b_0 = 0.0715$ ) in Experiment 3.

The results of regression analysis of the relationship between the IPLR for SIP signaling flows and the network/link load for each scenario of conducted experiments are summarized in Table 5. The cubic regression model is used to describe how the typical value of IPLR for SIP signaling flows changes when network/link load is varied in

Scenarios 1 and 2 of each experiment, as well as in Scenario 3 of Experiment 2. The very strong positive correlation is observed between the IPLR for SIP signaling flows and network/link load, which is statistically significant ( $P < 0.05$ ). The regression model of the relationship between the IPLR for SIP signaling flows and the network/link load cannot be determined for Scenario 3 of Experiments 1 and 3, as well as for Scenario 4 of each experiment, because the IPLR is kept constant regardless the network/link load. It means that there is no statistical relationship between the IPLR and the network/link load when SIP signaling flows are marked with the highest priority.

## 5. Conclusion and Future Work

The evolution of NGN raises the issue of exploiting customizable service for nomadic user within the environment of heterogeneity and mobility. In this context, the end-to-end QoS remains always as an issue. The novel approach based on prioritizing signaling information transmission is proposed to address this issue. It has been recognized that prioritizing the signaling information transmission is the main precondition to guarantee the service reachability and implement dynamic end-to-end QoS control depending on the availability of the current network resources and thereby improve transport network resource utility effectively.

This paper extends our previous work in terms of statistical analysis of signaling performance metrics. It shows that there is statistically significant relationship between the signaling performance metrics and network/link load when standard approach to signaling information transmission is used. Although the network links seem to be underutilized, the signaling performance degradation present in this case could be explained by microcongestion events. The detailed analysis of the frequency and duration of microcongestion is needed to determine its impact on the performance degradation on underutilized links. The results obtained in this work indicate that there is almost no statistical relationship between the signaling performance metrics and the network/link load when novel approach is used. This is important for guaranteeing service reachability and continuity throughout the session while moving or changing terminal.

In addition to summarizing our research findings, the opportunity for the improvement of the novel approach may be identified. Prioritizing signaling information transmission may be based on the network utilization. With the network utilization around the 90%, higher priority level should be assigned to those signaling packets that terminate session and, thus, reduce the overall network congestion. On the other hand, with the network utilization around the 20%, higher priority level should be assigned to those signaling packets that establish new sessions.

Two additional issues may be identified for future work. They are related to the difficulties in cross-domain and cross-layer priority level configuration for Signaling service class. The target QoS objectives for Signaling service class are ideally achieved by identical QoS marking and packet treatment policies across layers and networks involved in

the end-to-end packet transport. However, the local policy enforcement and possible remarking options likely lead to suboptimal signaling packet treatment traveling through several domains. Furthermore, the cross-layer QoS mapping occurs internally in an uncoordinated manner and is not signaled across network boundaries.

## Acknowledgment

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions, which significantly improved the quality of this paper.

## References

- [1] R. Stankiewicz, P. Cholda, and A. Jajszczyk, "QoX: what is it really?" *IEEE Communications Magazine*, vol. 49, no. 4, pp. 148–158, 2011.
- [2] J. Song, M. Y. Chang, S. S. Lee, and J. Joung, "Overview of ITU-T NGN QoS control," *IEEE Communications Magazine*, vol. 45, no. 9, pp. 116–123, 2007.
- [3] F. Weber, W. Fuhrmann, U. Trick, U. Bleimann, and B. Ghita, "QoS in SIP-based NGN—state of the art and new requirements," in *Proceedings of the 3rd Collaborative Research Symposium on Security, E-Learning, Internet and Networking (SEIN '07)*, pp. 210–214, University of Plymouth, June 2007.
- [4] J. Baraković, H. Bajrić, and S. Baraković, "Priority level configuration for signaling service class," in *Proceedings of the 3rd International Conference on Communication Theory, Reliability, and Quality of Service (CTRQ 2010)*, pp. 122–127, IEEE Press, Athens, Greece, June 2010.
- [5] J. Rosenberg, H. Schulzrinne, G. Camarillo et al., "SIP: session initiation protocol," Tech. Rep. RFC 3261, Internet Engineering Task Force (IETF), June 2002.
- [6] J. Baraković, H. Bajrić, and M. Kos, "Priority transmission of SIP signaling flows over IP network," in *Proceedings of International Conference on Multimedia Computing and Systems (ICMCS '09)*, pp. 96–101, IEEE Press, Ouarzazate, Morocco, April 2009.
- [7] J. Baraković, H. Bajrić, and M. Kos, "Priority transmission of sip signaling flows in case of ip link congestion," in *Proceedings of the 2nd International Conference on Communication Theory, Reliability, and Quality of Service (CTRQ '09)*, pp. 26–31, IEEE Press, July 2009.
- [8] J. Baraković, H. Bajrić, and M. Kos, "Priority transmission of SIP signaling flows in case of link failure," in *Proceedings of the 10th International Conference on Telecommunications (ConTEL '09)*, pp. 389–396, IEEE Press, June 2009.
- [9] S. Blake, D. Black, M. Carlson et al., "An architecture for differentiated service," Tech. Rep. RFC 2475, Internet Engineering Task Force (IETF), December 1998.
- [10] K. Nichols, S. Blake, F. Baker, and D. Black, "Definition of the differentiated service field (DS Field) in the Ipv4 and Ipv6 headers," Tech. Rep. RFC 2474, Internet Engineering Task Force (IETF), December 1998.
- [11] J. Baraković, H. Bajrić, and A. Husić, "QoS design issues and traffic engineering in next generation IP/MPLS network," in *Proceedings of the 9th International Conference on Telecommunications (ConTEL '07)*, pp. 203–209, IEEE Press, June 2007.
- [12] H. Hexian and D. Dongfeng, "Resource and admission control in NGN," *ZTE Communications*, vol. 5, no. 1, pp. 17–20, 2007.

- [13] S. Jun, "QoS issues in converged networks," *ZTE Communications*, vol. 5, no. 2, pp. 34–37, 2007.
- [14] J. M. Batalla and R. Janowski, "Provisioning dedicated class of service for reliable transfer of signaling traffic," in *Proceedings of the 20th International Teletraffic Congress on Managing Traffic Performance in Converged Networks (ITC20 '07)*, vol. 4516 of *Lecture Notes in Computer Science*, pp. 853–864, Springer, Ottawa, Canada, June 2007.
- [15] X. Xiao, *Technical, Commercial and Regulatory Challenges of QoS: An Internet Service Model Perspective*, Morgan Kaufmann, San Francisco, Calif, USA, 2008.
- [16] K. Papagiannaki, D. Veitch, and N. Hohn, "Origins of microcongestion in an access router," in *Proceedings of the 5th International Workshop on Passive and Active Network Measurement (PAM '04)*, vol. 3015 of *Lecture Notes in Computer Science*, pp. 126–136, Springer, Antibes Juan-les-Pins, France, April 2004.
- [17] J. Babiarz, K. Chan, and F. Baker, "Configuration guidelines for diffServ service classes," Tech. Rep. RFC 4594, Internet Engineering Task Force (IETF), August 2006.
- [18] V. Joseph and B. Chapman, *Deploying QoS for Cisco IP and Next Generation Networks: The Definitive Guide*, Morgan Kaufmann, San Francisco, Calif, USA, 2009.
- [19] F. Baker, J. Polk, and M. Dolly, "A differentiated services code point (DSCP) for capacity-admitted traffi," Tech. Rep. RFC 5865, Internet Engineering Task Force (IETF), May 2010.
- [20] "The Network Simulator - (ns-2), version 2," February 2011, <http://www.isi.edu/nsnam/ns/>.
- [21] Rui Prior's webpage, "ns-2 network simulator extensions," February 2011, <http://www.dcc.fc.up.pt/~rprior/ns/>.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

