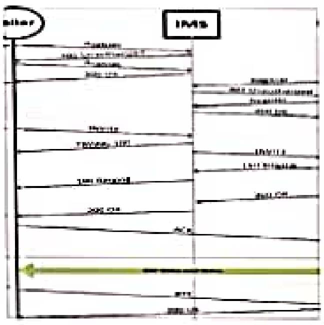
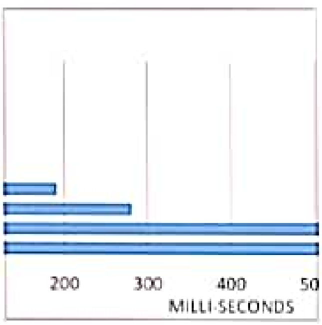
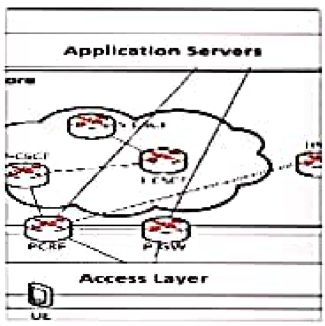
Mission-critical Communication Systems that are adaptable for use with the latest generation of multimedia services are crucial for system users. To determine the set of requirements that need to be hardcoded into such systems, a clear distinction between mission-critical and non-mission-critical systems is required. Moreover, the users of services provided by such systems are very different to those of current mobile commercial communication systems. These differences give rise to a set of challenges that need addressing to facilitate migration from existing systems to those now being proposed. One such challenge relates to the performance of the IP Multimedia Subsystem (IMS) registration process. This is a crucial consideration for mission-critical systems, particularly in large-scale systems where thousands or even millions of users may seek to access the system in disaster scenarios. This paper presents an evaluation of IMS and Session Initiation Protocol (SIP) performance metrics and Key Performance Indicators (KPIs). Moreover, it articulates a proposed study that will seek to address some of the challenges identified.



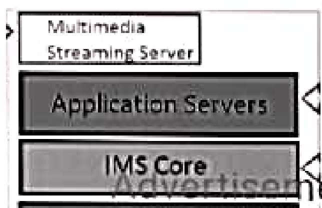
SIP signalling flow and...

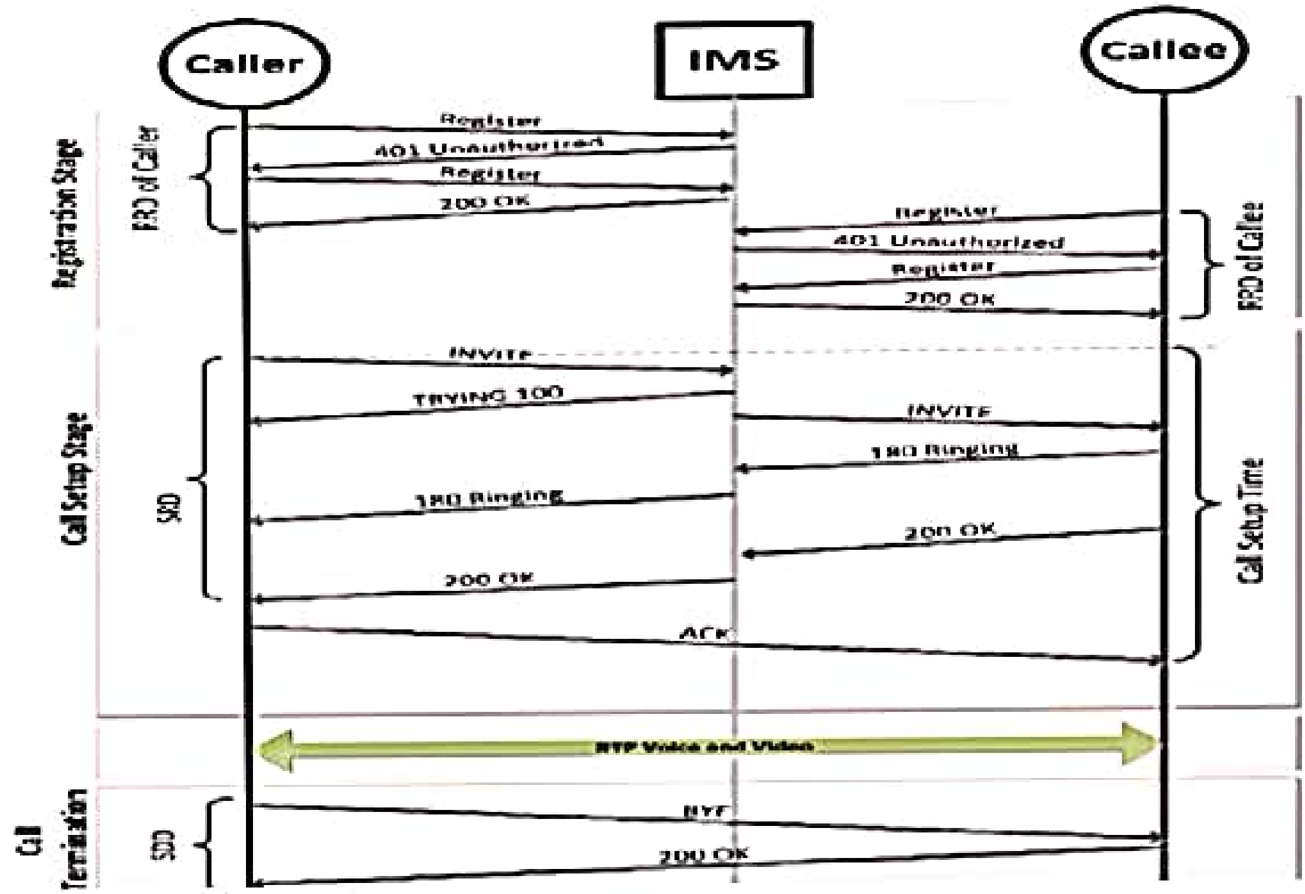


Access Time Evolution for...



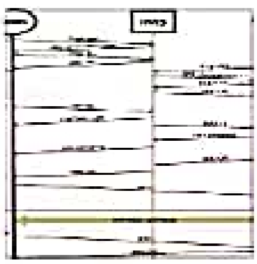
IMS Charging Functions

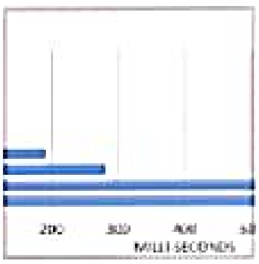


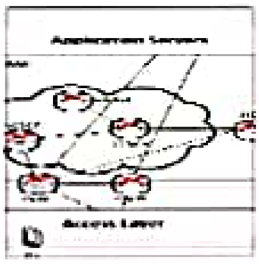


SIP signalling flow and performance metrics.

Source publication







Effective Performance Metrics for Multimedia Mission-critical Communication Systems

Article

Full-text available

'Mission Critical' is defined by TETRA and The Critical Communication Association (TCCA) as a function whose failure would lead to catastrophic consequences that would place public order or public security at risk. A system that provides such critical functionality must have suitable inbuilt functionality, interoperability, security, and the wherewithal to maintain its availability. Mission-critical users are those with responsibility for the welfare, health, security and safety of the public. Law enforcement forces, firefighters, emergency and medical services, rescue services, military forces, utility staff members, and transport services members, are all example of mission-critical users. The concept of a Mission-critical Communication Systems (MCCS) refers to the hardware, software and communication facilities, that allow mission-critical users to communicate with each other and liaise with command centres securely and dependably for the sake of providing mission-critical services, wherever and whenever the services require special communication solutions.

Many governmental and non-governmental organisations (NGOs) play critical roles in the operation of mission-critical systems. The services provided by such organisation are of a type that makes them less tolerant to execution or operational errors. Such services are referred to as mission-critical services. The criticality of such services implies that they have a set of special requirements that distinguish them from other services. They should be available anytime, anywhere, within the

Ashraf Ali and Andrew Ware, "Effective Performance Metrics for Multimedia Mission-critical Communication Systems", *Annals of Emerging Technologies in Computing (AETiC)*, Print ISSN: 2516-0281, Online ISSN: 2516-029X, pp. 1-14, Vol. 5, No. 2, 1st April 2021, Published by International Association of Educators and Researchers (IAER), DOI: 10.33166/AETiC.2021.02.001, Available: http://aetic.theiaer.org/archive/v5/v5n2/p1.html

service operational scope. Moreover, they need to be able to function at the extremity of their capabilities regardless of the operational circumstances or running conditions.

The most important protocol performance issues that control the operation of end-to-end systems and have implications on its overall performance are presented in this paper. There are mainly two types of protocols: control plane protocols and data plane protocols. The Session Initiation Protocol (SIP) is one of the control plane protocols and was fully standardised by the Internet Engineering Task Force (IETF) in RFC 2543 for the first version of SIP 1.0 and RFC 3261 for the second version of SIP 2.0 [1]. The SIP is a communication protocol that operates over IP and is used for the signalling of real-time multimedia services such as voice, video and non-real-time services (for example, text messages and presence notifications). The text-based protocol mainly defines the signalling order between end-users for call initiation and termination, in addition to instantly modifying the call setup as needed. It is also used for registering the users before the call is initiated.

Although the SIP message headers and signalling details are not presented in this paper, performance issues and the challenge of enhancing SIP services performance are briefly discussed.

2. SIP Performance Issues

The protocol related performance metrics need to be identified to determine how the SIP is utilising the system resources and to help ensure optimal usage. Moreover, the architectural design challenges need to be targeted to enhance the SIP performance. A variety of tests that measure the processing time of SIP messages, memory allocation, thread performance, and call-setup time was presented [2]. The results show that the performance of the proxy server is affected by varying SIP related parameters, thus affecting the number of calls per second that the proxy server can handle. The network topology, such Wireless Sensor Networks (WSNs) affect the SIP performance requirements [3]. The results also show that the performance of SIP related architectures, such as the IP Multimedia Subsystems (IMS) presented later, are more affected than a simple Proxy/Registrar Server due to the heavy dependence of SIP signalling and messages structure. Furthermore, it is important to note that the performance of SIP signalling is significantly affected by the delay at different stages of registration, call initiation, and call termination processes. The performance of SIP signalling also affects the QoS of the offered service. Hence, the need to define the metrics that identify the performance measure for the SIP is crucial for evaluation and performance comparison purposes. The IETF proposed the criteria for the end-to-end SIP performance measures in RFC 6076 [4].

Understanding SIP performance is essential for determining the overall multimedia system communication performance. IMS, for example, is composed of many entities that communicate with each other using SIP signalling. In other research, the impact of the scheduler settings on the overall performance of multithreaded servers was investigated [5]. The work showed that the operating system and the hardware architecture of the machines that host the SIP server have a significant impact on the SIP performance. Service scheduling for multithreaded systems, for example, is controlled via the process scheduler that ensures that the multi-process and threads within the SIP registration process successfully. SRD is the time required to get a reply from the server-side regarding the requested call setup from the user side; it is counted for both successful and unsuccessful call requests. If the call requested were successfully set then the call setup time would simply be the SSD plus the acknowledgement sending time. The SDD is the time difference between sending the BYE message from the user side and the time of receiving 200OK confirmation from the server.

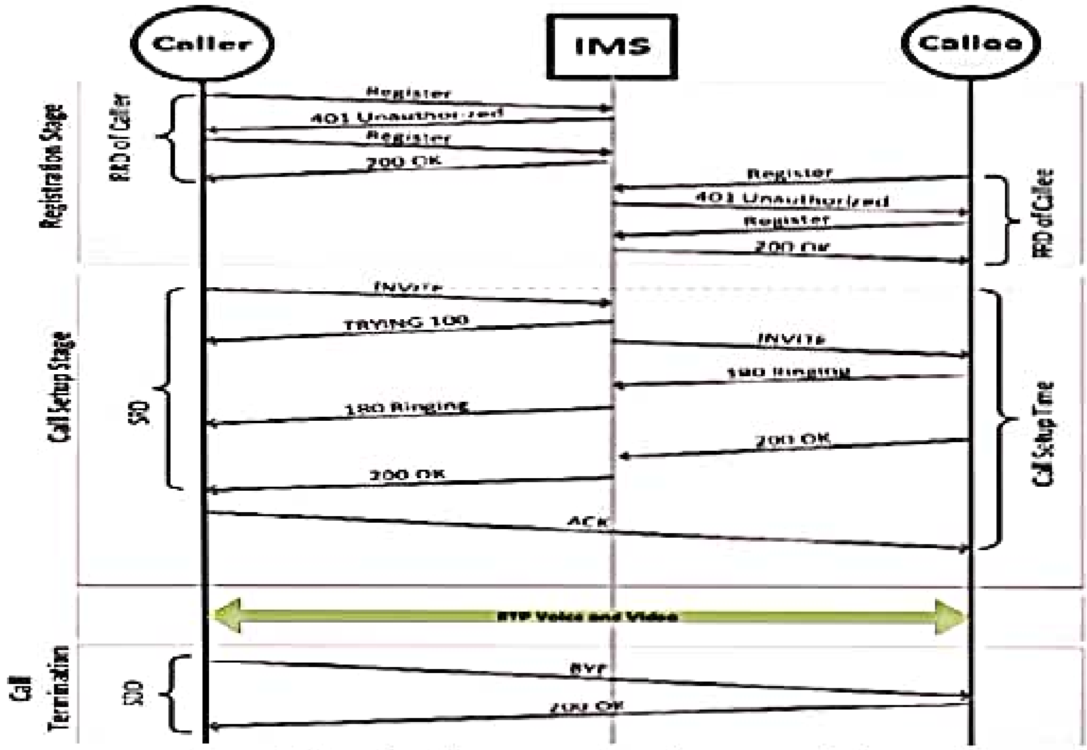


Figure 1. SIP signalling flow and performance metrics.

In this paper, the call setup delay is used to measure the system performance due to its importance in real-time multimedia services in general and in mission-critical communications specifically. The QoE for SIP-based systems is affected significantly by the call setup value. The mean time of call setup values can reach up to 800 ms [6]). However, for LTE-based mission-critical systems (designed to work as a replica for traditional dedicated mission-critical systems such as TETRA), the average call setup time needs to be within 500 ms delay. Figure 2 shows the average access time of different technologies. The call setup time is usually a multiple of the access time due to the enforced handshaking process.

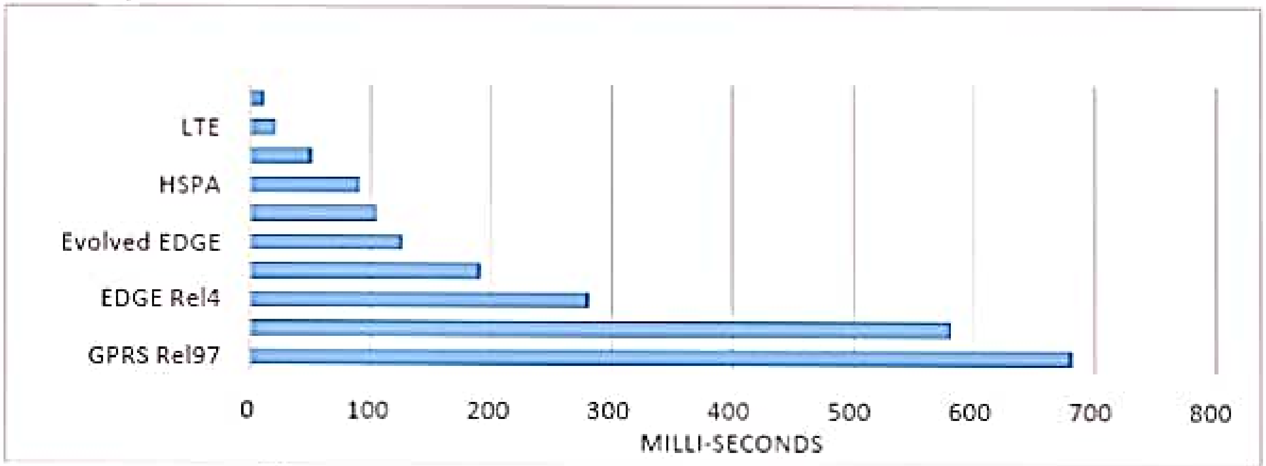


Figure 2. Access Time Evolution for Different Technologies

3. IP Multimedia Subsystem (IMS)

Legacy mobile communication networks and the next generation systems can provide end-users with new sets of applications and services. There is a need for a system that sits on top of the access technology domain to provide the required signalling for multimedia services. IP Multimedia Subsystems (IMS) form the core part of Next Generation Networks (NGNs). They play an important role in helping the service operators merge the multimedia services in cellular networks and to provide end-users with key features and services within Quality of Service (QoS) levels set by operators. These multimedia services such as messaging, instant voice, video conferencing, group management, and push services, all rely on IMS to control the signalling and the state of the service before initiation, during the service and after service termination. IMS can be integrated in non-multimedia based systems as well [7].

IMS was designed and standardised by 3GPP to provide multimedia services over mobile communication technologies beyond GSM [8]. The IMS is used for delivering the IP multimedia services between users and service providers; it gains its importance as an architectural framework for multimedia communications, it can also be integrated in other frameworks for robot messaging and logging [9]. The SIP is the principal signalling protocol used within an IMS to create, modify and destroy multimedia sessions. As defined by the standard, an IMS functions as an interface between the service/application layer and the transport layer, enabling the service providers and operator to control the user QoS based on its subscription profile. Moreover, it works as a hub point for the entire SIP signalling that needs to take place before, during, and after the call. To ensure overall integrity, the different control functions are connected through a set of interfaces.

Users' subscription-related information is stored in the Home Subscriber Server (HSS) that performs authentication and authorisation functions for users. The HSS is also responsible for updating user registration status records. The Call Session Control Functions (CSCF) are responsible for handling the SIP signalling messages and packets in the IMS. The Proxy CSCF (P-CSCF) is the entry point into the IMS system, and all SIP messages flow through the P-CSCF. The P-CSCF may also apply security or compression algorithms to the received traffic in addition to providing quality of service control and bandwidth management. The interrogating CSCF (I-CSCF) is one of the main elements of the IMS systems. It is used during the registration process when the UE does not know which Serving CSCF (S-CSCF) should receive the request. The I-CSCF interrogates the HSS to obtain the address of the appropriate S-CSCF that should process the request.

The S-CSCF performs session control functions that interface with the HSS to check and download the user profile information. It assigns the Application Server (AS) for the user for further services and enforces the operator policy control. The SIP AS has a SIP interface with the S-CSCF and is used to host specific IMS services. After the registration process is completed and the S-CSCF is allocated to the UE, the I-CSCF is no longer used for any further communication. All future communication happens between the UE, the P-CSCF, and the S-CSCF.

In the next sections, the performance of IMS, which is critical in affecting the overall system performance due to its hierarchal position in the stack and core functional role, is discussed.

3.1. IMS Performance Issues

After migrating from the circuit-switched Second Generation Mobile Networks (2G) toward the packet-switched domain of fourth-generation (4G) communication networks and beyond, the need for supporting multimedia services in all IP network infrastructure is essential to ensure the convergence of data, multimedia services, and mobile networks technologies. IP Multimedia Subsystems (IMS) developed by the Third Generation Partnership Project are a key part of Next Generation Networks (NGN) that are responsible for providing and controlling the multimedia services in the packet-switched domains. As defined by the standard, the IMS operates as an interface between the service/application layer and the transport layer that enables the service providers and operator to control the user QoS based on its subscription profile. Moreover, it works as a hub point for the entire SIP signalling that needs to take place before, during, and after the call. For this purpose, there are different functions that are connected by interfaces to ensure operational integrity.

In the protocol stack, the IMS resides between the application layer and the transport layer (see figure 3). Therefore, the performance of the IMS affect mission-critical applications that have strict requirements due to their special operating nature and associated tasks. End-end delay of the access technology components along with the IMS entities affect the end-to-end signalling, and QoE offered to the end-user. Having a generic system that has IMS core network serving both legacy mobile users

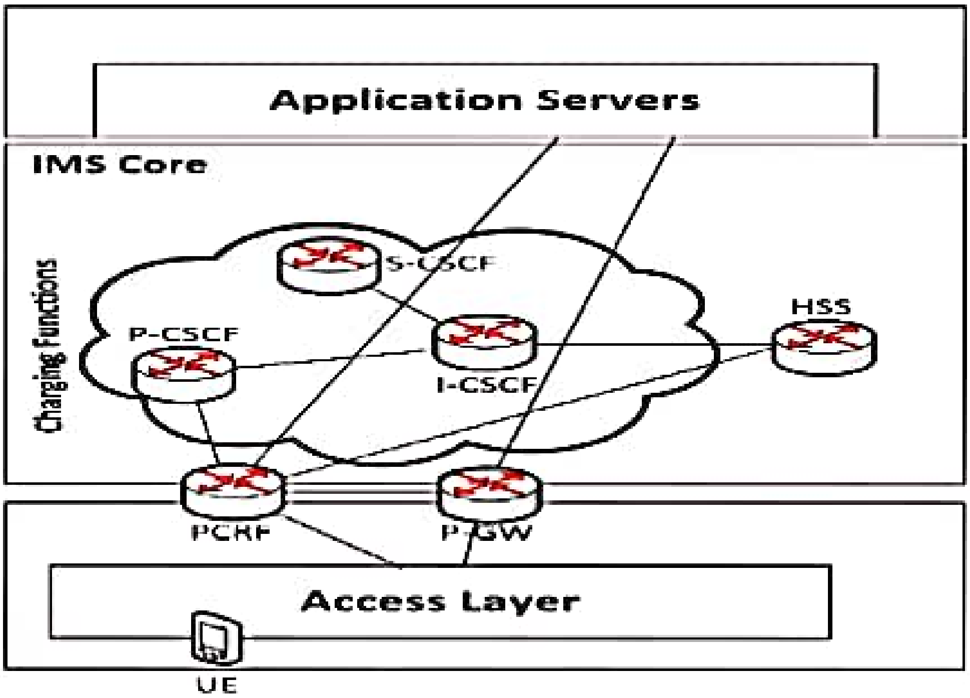


Figure 3. IMS Charging Functions

The performance of the IMS is affected by the individual performance of each entity inside it. The P-CSCF is the first entity that interacts with the User Element (UE) and forwards the SIP messages to other control functions in the IMS. It is also used to apply security or compression algorithms to the received traffic in addition to monitoring quality of service and bandwidth management. The I-CSCF is a SIP server that is assigned by the HSS to the user when the I-CSCF requires it. The S-CSCF has interfaces with the HSS to perform session control, download the user profile information, assign the Application Server (AS) for further services, and to enforce the operator policy control. The SIP AS has a SIP interface with the S-CSCF, and is used to host specific IMS services. Some gateways function as interfaces with other domains.

3.2. IMS Registration Performance

In the IMS registration, the user requests authorisation to gain access to the IMS services. Figure 4 shows the registration steps in IMS. First, the UE sends a SIP Register request to the P-CSCF, which forwards it to the I-CSCF. Then the I-CSCF sends a User Authentication Request (UAR) to the Home Subscriber Station (HSS), using the Diameter protocol, to check the user profile for authorisation process and to determine the S-CSCF address allocated to the user. Then, the HSS replies with the User Authentication Answer (UAA) to the I-CSCF and authorises the user. After retrieving the S-CSCF details from the HSS, the I-CSCF forwards the SIP Registration request to the S-CSCF, which in turn sends a Multimedia Authentication Request (MAR) to download user authentication data to the HSS. The HSS then replies with the Multimedia Authentication Answer (MAA) to the S-CSCF, which in turn responds to the user using a SIP 401 Unauthorised response that embed an authentication challenge within it for the UE. The UE then generates another SIP Register request following the same steps as described earlier. This time the authentication process finishes and the S-CSCF sends a Server Assignment Request (SAR) to the HSS, using the Diameter protocol, which replies with a Server Assignment Answer (SAA) using the same protocol. Finally, the S-CSCF sends a SIP 200OK message to the UE to complete the registration process.

As mentioned before, the Registration Request Delay (RRD) is the time needed for the registration process to complete. In this case, that is the time difference between the first SIP Register message sent by the UE and the time the SIP 200 OK message is received. Figure 5 shows the transition diagram of the IMS Registration process.

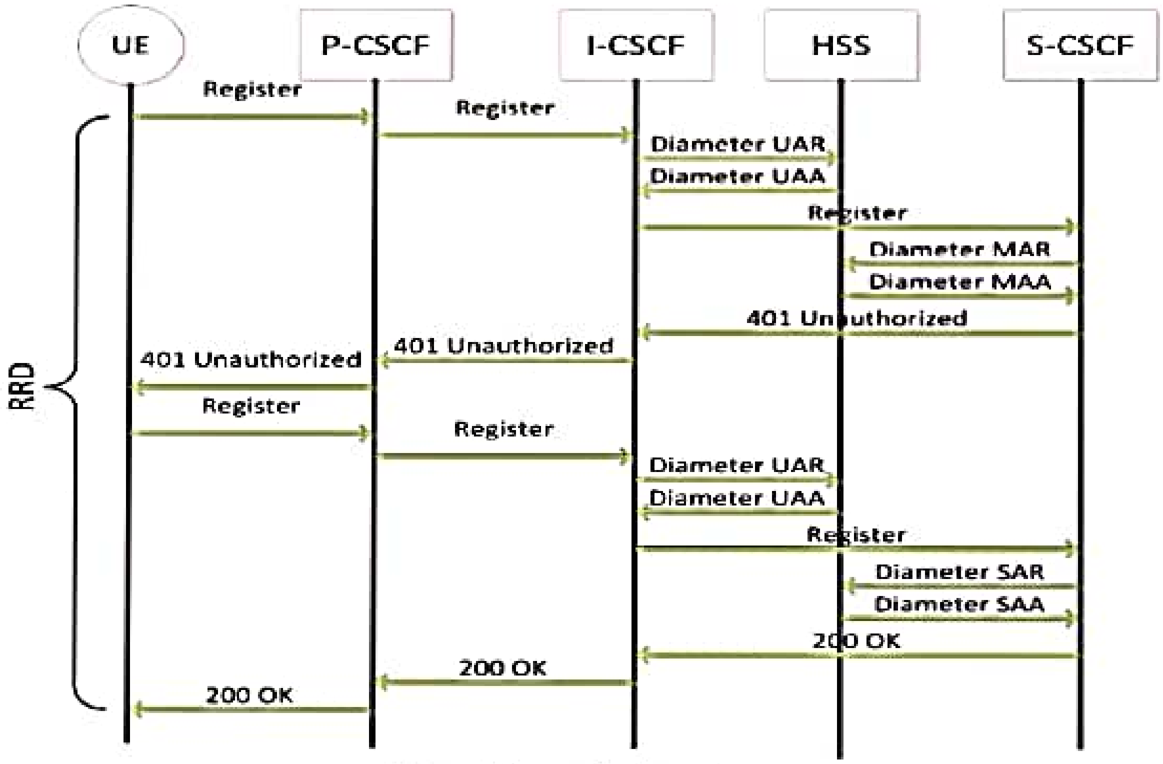


Figure 4. IMS Registration Process

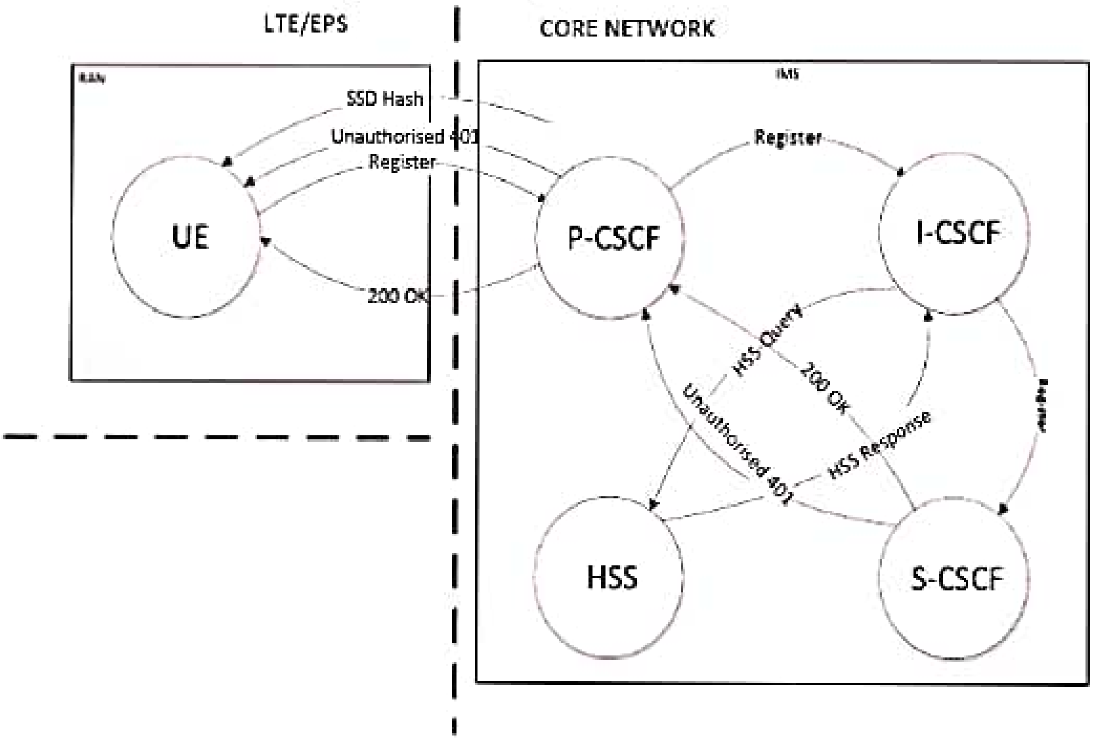


Figure 5. Transition Diagram of IMS Registration Process

During the registration process, the end-user needs to be subscribed to the IMS to be authorised to make calls and initiate other services instantly during the call. The registration process is required before starting calls (see figure 6) and reregistration is needed for the SIP device to re-authenticate itself after receiving a notification from the SIP server. This is required to avoid compromising the SIP device.

The delay in the registration process is one of the SIP performance metrics and criteria that was defined by IETF in RFC 6076 [4], due to the previous lack of a SIP benchmarking criteria to define the baseline performance of SIP signalling. RFC 6076 defines the performance metrics for the SIP in VoIP applications in order to provide key performance indicators and Service Level Agreement (SLA) indicators for best network resources utilisation and best end-user Quality of Experience (QoE). The main metrics defined in RFC 6076 are the Registration Request Delay (RRD), Ineffective Registration Attempts (IRA), Session Request Delay (SRD), and Session Disconnect Delay (SDD). RRD is the time needed for the user to finish the registration process successfully and plays a major role in the overall QoS.