3GPP TR 23.820 V9.0.0 (2009-09)

Technical Report

3rd Generation Partnership Project; Technical Specification Group Core Network and Terminals; Study on IMS Restoration Procedures (Release 9)

Keywords
UMTS, LTE, GSM, IP, multimedia, network

3GPP

Postal address

3GPP support office address
650 Route des Lucioles - Sophia Antipolis
Valbonne - FRANCE
Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Internet
http://www.3gpp.org

Contents

Foreword

This Technical Report has been produced by the 3rd Generation Partnership Project (3GPP).
The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

- x the first digit:
  - 1 presented to TSG for information;
  - 2 presented to TSG for approval;
  - 3 or greater indicates TSG approved document under change control.
- y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.
- z the third digit is incremented when editorial only changes have been incorporated in the document.

Introduction

Although network nodes in the IMS Core Network should have a very high availability, some maintenance downtime and occasional failures are unavoidable. Communication links although designed with robust protocols between the network elements are also subject to failures. A set of standardized procedures for automatic restoration after loss or corruption of data could reduce the impact of these problems resulting in the improved service to the users. The intention is that similar cases as in 3GPP TS 23.007 [2] for the CS and PS Domains are covered also for the IMS domain.

1 Scope

The present document identifies the changes required in the 3GPP IMS specifications so that a consistent state is restored in the IMS Core Network, after, or during a planned, or unplanned stop of a network element. The study will go through the following steps:

- Establish the requirements that should be covered with these procedures. That is which are the impacts to the end user service that are acceptable and which are not, after a network failure.
- List the service interruption scenarios that need to be studied.
- Provide solutions, so that in all the service interruption scenarios listed, the impacts to the end user service comply with the requirements. These solutions provide procedures for the automatic restoration to a consistent state in the network and indicate how to trigger these procedures.
- Analyze the impacts of the solutions in the current specifications.
- Conclusion and recommended way forward.

It is important to realise that these procedures are meant to be operational procedures for restoration and so care must be taken with what is existing and will exist with OA&M procedures to avoid overlap which could cause clashes.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".
[2] 3GPP TS 23.007: "Restoration procedures".
[3] 3GPP TS 32.260: "Charging Management; IP Multimedia Subsystem (IMS) Charging".
[4] 3GPP TS 23.228: "IP Multimedia (IM) Subsystem – Stage 2".
[5] 3GPP TS 24.229: "IP Multimedia Call Control Protocol based on SIP and SDP".
[6] 3GPP TS 23.008: "Organization of subscriber data".
[7] IETF Draft draft-ietf-sip-outbound: "Managing Client Initiated Connections in the Session Initiation Protocol (SIP)".
[8] 3GPP TS 29.061: "Interworking between the Public Land Mobile Network (PLMN) supporting packet based services and Packet Data Networks (PDN)".
[9] 3GPP TS 29.060: "General Packet Radio Service (GPRS); GPRS Tunnelling Protocol (GTP) across the Gn and Gp interface".
[10] 3GPP TS 29.228: "IP Multimedia (IM) Subsystem Cx and Dx interfaces; Signalling flows and message contents".
[11] IETF RFC 3588: "Diameter Base Protocol".

3 Definitions, symbols and abbreviations

3.1 Definitions

For the purposes of the present document, the terms and definitions given in TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in TR 21.905 [1].

Service Interruption: A period of time in which one or more network elements do not respond to requests and do not send any requests to the rest of the system.

3.2 Abbreviations

For the purposes of the present document, the abbreviations given in TR 21.905 [1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in TR 21.905 [1].

GRUU	Globally Routed User Agent URI
LIA	Location-Info-Answer
LIR	Location-Info-Request
MAA	Multimedia-Auth-Answer
MAR	Multimedia-Auth-Request
OA&M	Operations, Administration & Maintenance
O&M	Operations & Maintenance
OPEX	Operative Expenditures
SAA	Server-Assignment-Answer
SAR	Server-Assignment-Request
UAA	User-Authorization-Answer
UAR	User-Authorization-Request

4 Requirements analysis and assumptions

4.1 Introduction

This clause contains a list of the requirements for the solution provided in this study. The general goal is to have a set of procedures ensuring that the impact of the service interruption of a node is limited to the loss of the capacity of that node for the time that it is out of service, plus some additional signalling in order to perform the take over by other network elements with the same function.

4.2 Persistency Requirements for Data

There are network elements that hold permanent data. There should be methods to ensure that the information in these network elements is not lost even in disaster events, such as a complete site crash. For this reason, this study assumes that the redundancy provided for these network elements makes unnecessary additional network procedures for restoration of permanent data. On the other hand, temporary data will be considered in these procedures.
For the nodes that don't handle permanent data, the assumption is that their memory is affected if an outage occurs and the information related to some of the users may be lost.

4.3 Impacts on Established Sessions

Interruption of established sessions is considered an acceptable consequence of the failure service interruption of one of the network elements in the session path. This means that the restoration of session data does not need to be analyzed in this study. Means may be taken by implementations outside of this study to enable established sessions to be restored or maintained.
The accounting of IMS sessions is already based on principles (interim accounting as described in 3GPP TS 32.260 [3]) that ensure that the charging of this interrupted sessions is also terminated.

4.4 Impacts on Session Establishment Time

The restoration procedures could involve some steps that take place during the establishment of sessions after the outage has occurred (i.e. after the network element that failed returns to normal functioning or another network element takes over). If that is the case, the increase of the session establishment time should not be significant (it should still have a high probability to remain within the acceptable levels for the end user).

4.5 Required Manual Intervention

In order to reduce OPEX and the time required to restore the network, the need for manual intervention should be minimized. This implies that the procedures should be triggered by network signalling events and O&M steps should also be minimized.

4.6 Loss of service

Loss of service refers to the state in which session origination attempts by the user or session termination requests to that user fail while the UE appears to be registered and also when the network does not respond to registration attempts. Ideally the proposed solution should avoid this kind of loss of service altogether and ensure that requests are terminated correctly in all cases. If this is not feasible, then the time of loss of service for the user should be minimized.

4.7 Avoidance of massive signalling

In the solutions provided it needs to be taken into consideration that network element failures tend to occur in situations when the signalling is overloaded. If that is the case, restoration procedures that involve a high level of messages (e.g. triggering re-registrations for all the UEs controlled by a P-CSCF or S-CSCF) should be avoided. Such kind of procedures could result in further problems in other network elements and provoke a domino effect of subsequent failures.

4.8 Load balancing

The solution provided should be such that it allows the recovery of the network to a situation where the load is balanced between network elements performing the same function.

5 Service Interruption scenarios

5.1 Introduction

According to the requirements for this study, service interruption of network elements holding only persistent data or holding only session related data should not be considered. The following figure shows the network elements of the IMS Core Network. Those holding only persistent data are represented with cylinders and those holding only session data with round edges. According to this the targets for the study are coloured red in the following figure.

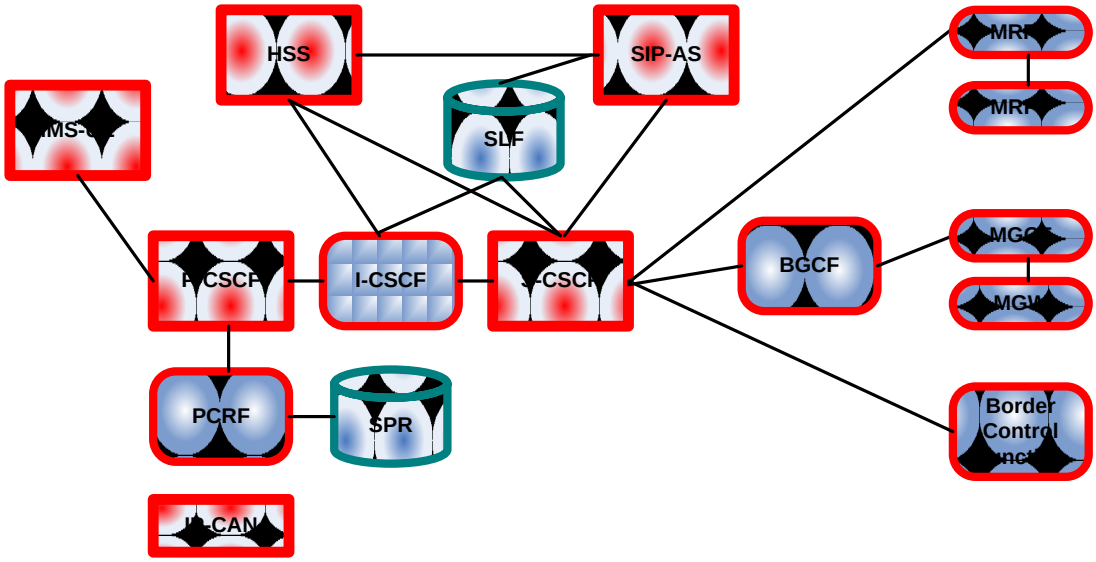

Figure 5.1: Network elements to be considered in the study

The following clauses cover each of the possible service interruption cases and indicate how they affect service.

5.2 S-CSCF Service Interruption

5.2.1 Introduction

This clause will analyse the impacts of a S-CSCF service interruption in the network and in the service to the user in order to highlight the problems that need to be covered by the alternative solutions in this technical report. The initial state that will be considered is an IMS Core Network working properly with several S-CSCFs and with ongoing traffic

(a certain amount of users are registered and unregistered in the S-CSCFs). At one point one S-CSCF stops operation, this may imply either a lack of response from that S-CSCF or potential loss of the information of some subscribers in that S-CSCF. The assumption will be that the S-CSCF does not trust any data after it resumes operation, due to the fact that it could have lost profile updates from the HSS in the service interruption period. The following clauses cover the outcome in several different cases.

5.2.2 Unregistered User

The lack of service of the S-CSCF in this case will mean that unregistered triggers for the user are not processed. Until the previously assigned S-CSCF resumes operation or the user initiates a registration the following occurs:

- The terminating SIP requests from the I-CSCF to the assigned S-CSCF will fail and there will be no service of unregistered services for the user.
- Originating request from a SIP-AS will also fail, since they will be forwarded to the assigned S-CSCF (directly, via the I-CSCF, etc.).

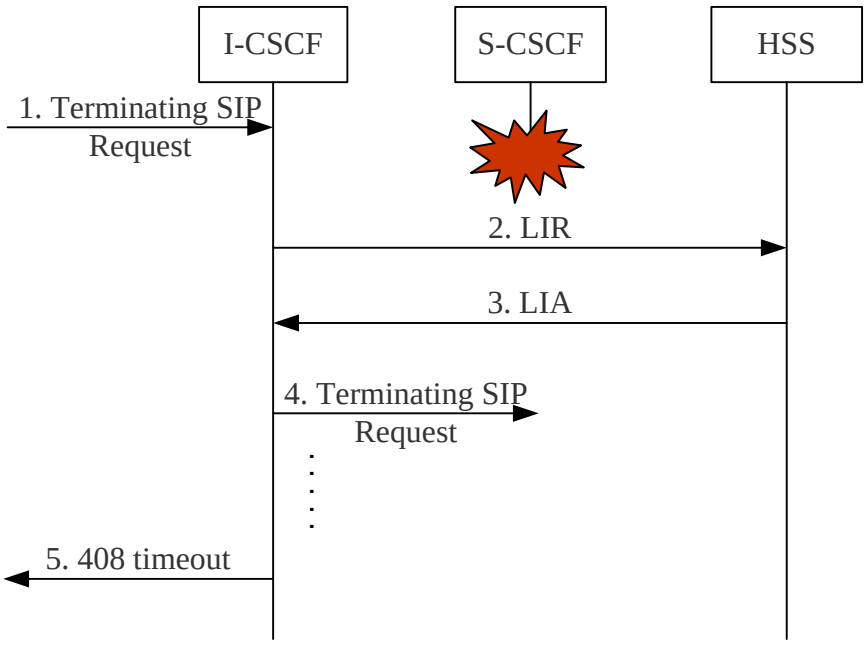

Figure 5.2.2.1. Terminating request to unregistered user with no S-CSCF response

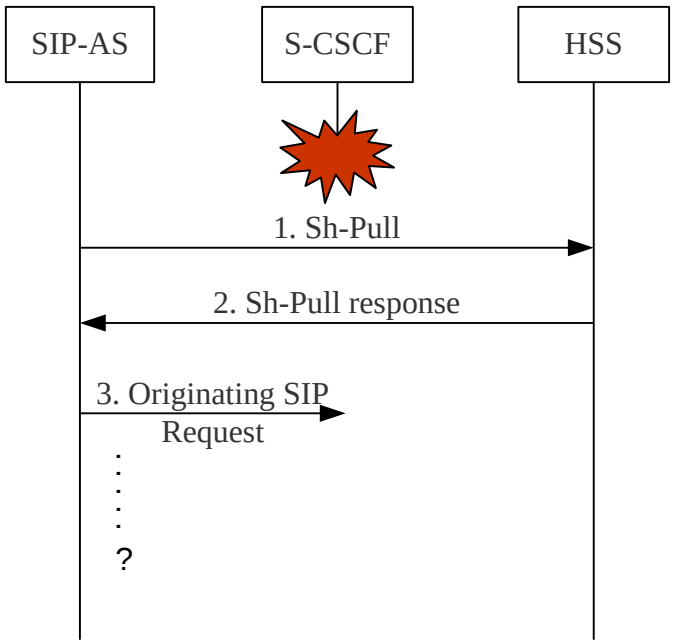

Figure 5.2.2.2. Originating request on behalf of unregistered user with no S-CSCF response

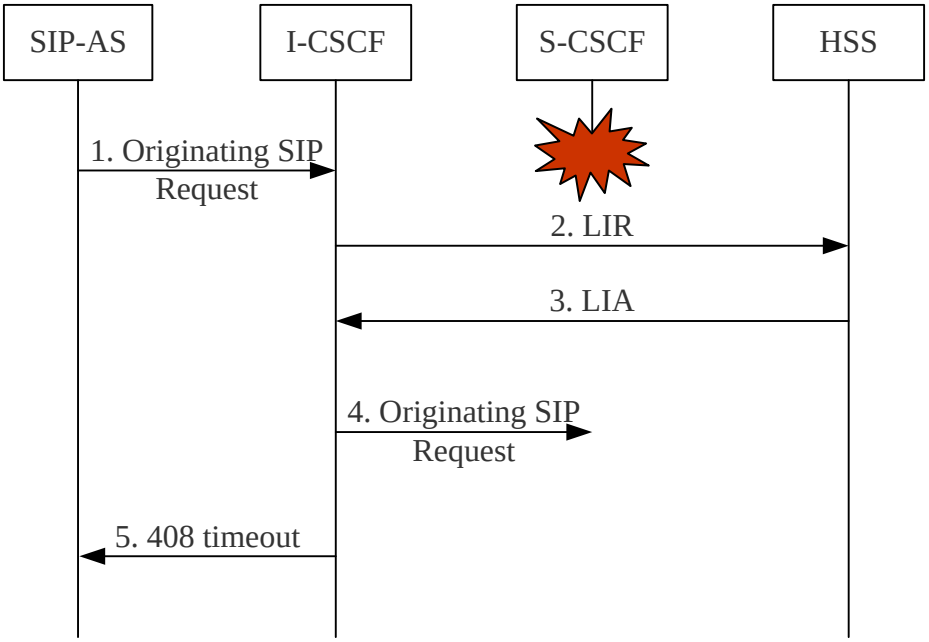

Figure 5.2.2.3 Originating request on behalf of unregistered user via I-CSCF

Even after there is a reassignment of S-CSCF (because of an initial registration) or the previously assigned S-CSCF resumes operation, there could be some problems with the handling of SIP subscriptions to notifications. This is because network elements could have subscribed to notifications in the previously assigned S-CSCF, and they will not be notified of the operation interruption or of the reassignment. The result is that these network elements will not receive the notifications that they expect until they subscribe again. Additional network procedures could reduce the unavailability of service in this scenario.

5.2.3 Originating Traffic from Registered User

In this case, the lack of response from the S-CSCF implies an error response to the UE that should trigger a new registration (depending on UE implementation). After the successful initial registration takes place, the initial request may be retried and normal operation will continue.

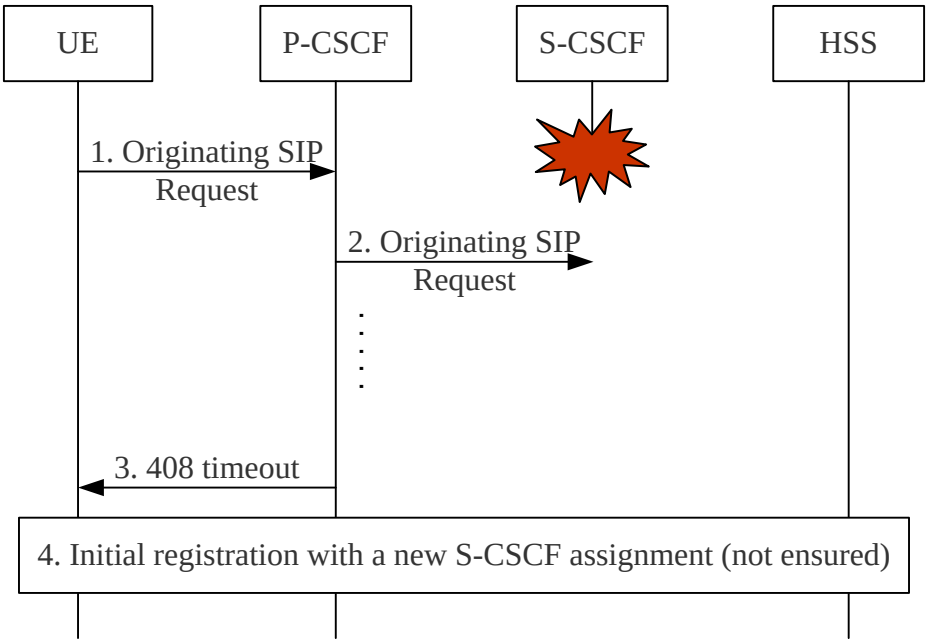

Figure 5.2.3.1. Originating request from registered user with no S-CSCF response

If the previously assigned S-CSCF resumes operation and it receives the request after losing the data for that user, it will reply with the corresponding SIP error, which should also trigger a new registration (depending on UE implementation).

However, SIP subscriptions to notifications in the previously assigned S-CSCF may be affected as indicated in clause 5.2.2.

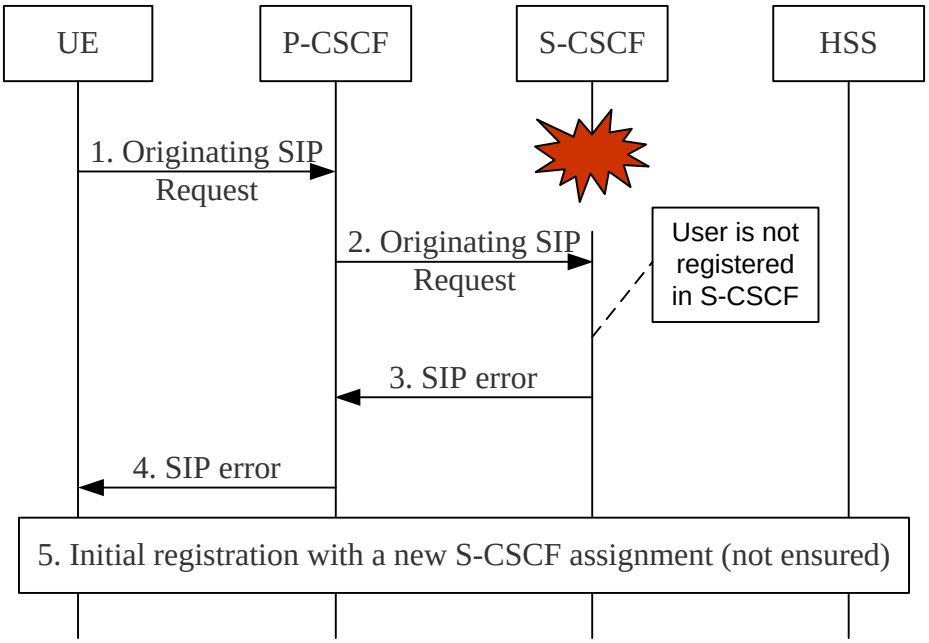

Figure 5.2.3.2. Originating request to registered user with loss of S-CSCF data

Traffic may also be originated on behalf of a registered user by a SIP-AS. The behaviour of the network in this case is the same, with the SIP-AS taking the role of the P-CSCF in the two previous flows.

5.2.4 Terminating Traffic for Registered User

It will not be possible to process correctly terminating traffic for a registered user until a new initial registration takes place. This situation may not be solved correctly if the previously assigned S-CSCF resumes operation. If the previously assigned S-CSCF loses the information related to one user, this will mean that after the S-CSCF returns to operation it will not identify this user as registered and it will send an SAR to the HSS requesting the change of state to "unregistered". The HSS will handle this as an error and the request will not be processed. This abnormal situation will persist until a re-registration takes place for the affected identities. The user will not be aware of this situation unless a SIP request is initiated from the UE. Considering that the re-registration timers can be quite long and also that the terminating user may not be initiating requests by itself regularly (in particular not during non-busy hours), the consequence of this can be seen as quite severe, and there is a need to try to improve the service availability in this scenario. SIP subscriptions to notifications in the previously assigned S-CSCF may also be affected as indicated in clause 5.2.2.

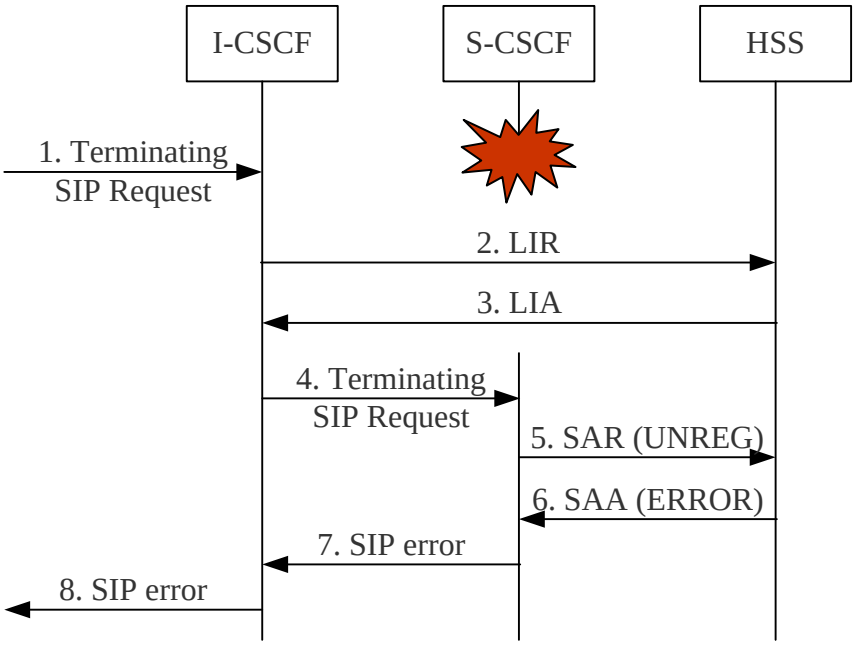

Figure 5.2.4.1. Terminating request to registered user with loss of S-CSCF data

5.3 P-CSCF Service Interruption

5.3.1 Introduction

This clause will analyse the impacts of a P-CSCF stop in the network and in the service to the user in order to highlight the problems that need to be covered by the alternative solutions in this technical report. The initial state that will be considered is an IMS Core Network working properly with several P-CSCFs and with ongoing traffic (a certain amount of equipments have security associations established with the P-CSCFs). At one point one P-CSCF stops operation, this implies lack of response from that P-CSCF and potential loss of the information of some subscribers in that P-CSCF.

5.3.2 Originating Traffic

In this case, the lack of response from the P-CSCF should trigger a new registration. After the successful initial registration takes place, the initial request may be retried and normal operation will continue. If the previously assigned P-CSCF resumes operation and it receives the request but it has lost the data for that user, it will ignore the request (the user appears as not authenticated), which should also trigger a new registration.

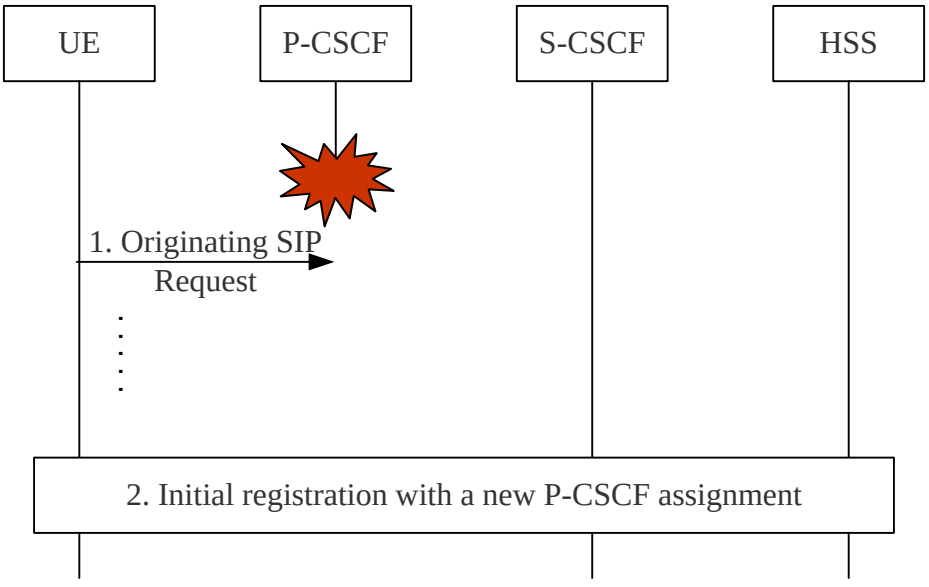

Figure 5.3.2.1. Originating request with no P-CSCF response

5.3.3 Terminating Traffic

In this case, the lack of response from the P-CSCF or error response from the P-CSCF because of loss of user information will mean that SIP requests will not proceed. The terminating requests to that user will fail and there will be no service for terminating requests until the user makes a new registration. The user will not be aware of this situation unless a SIP request is initiated from the UE. Considering that the re-registration timers can be quite long and also that the user may not be initiating requests by itself regularly (in particular not during non-busy hours), the consequence of this can be seen as quite severe, and there is a need to try to improve the service availability in this scenario.

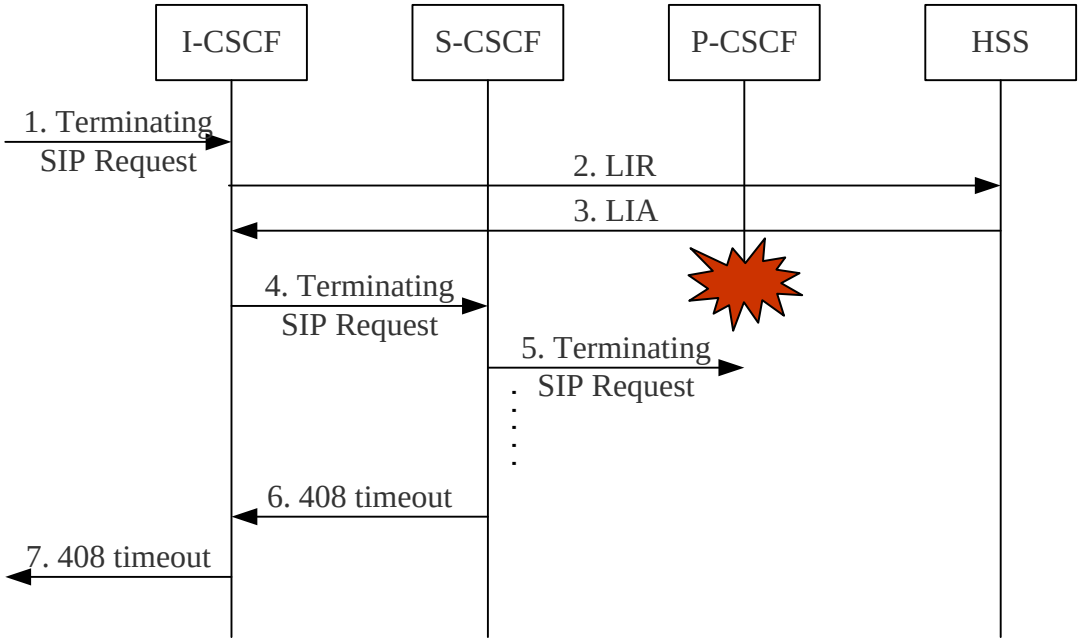

Figure 5.3.3.1. Terminating request with no P-CSCF response

5.3.4 Subsequent requests towards P-CSCF without traversing the S-CSCF

All SIP signalling to or from the UE traverses the P-CSCF, and all initial requests to or from the UE traverse the S-CSCF assigned to the UE. But according to TS 23.228 [4] section 5.4.5, it is not required that all subsequent requests traverse the S-CSCF. The section 5.4.5 and annex F of TS 23.228 describe this case in detail.
Some subsequent requests towards the P-CSCF will not traverse the S-CSCF, after the initial procedure. One case is when the subsequent requests are from SIP-AS. For example for the Presence service, notifications need not go through the S-CSCF. Another example could be where the UE initiates a session to an Application Server (AS) in the home operator's domain, e.g. video download.
In these cases, all subsequent requests towards the P-CSCF will fail, because of the failure of the P-CSCF.

5.4 IMS-UE Service Interruption

If the IMS-UE is not working this may imply total or partial lack of service for the user. However, the following reasons allow giving less importance to this kind of failure:

- There are already procedures to clean up the state of registration of these users in the network (the re-registration timer) and there also exists a handling today for incoming sessions when a UE is not reachable. The procedures describing how to handle an unreachable terminal are in clauses 5.10.3 and E.2.1a.2 of 3GPP TS 23.228 [4] and in clause 5.2.8.1 of 3GPP TS 24.229 [5].
- Handling interruptions in the UE and notifying the user is very implementation specific. Ultimately, the UE would make the user aware of the lack of service.

The conclusion is that solutions in this technical report do not need to add additional coverage for this case.

5.5 SIP-AS Service Interruption

5.5.1 Introduction

The SIP-AS may have stored transparent data in the HSS, and it may have subscriptions to notifications both using SIP and the Sh Interface. With respect to transparent data, when the SIP-AS returns to operation after a failure, it could be assumed that the same instances will be used again. It could also be assumed that the subscriptions to notifications are cleaned up periodically. Hence, the following discussion will only cover the case before the SIP-AS resumes. According to the section 4.3, interruption of established sessions is considered an acceptable consequence of the failure of one of the network elements in the session path, so restoration when the AS fails after it has been in the route path of an established session will not be analyzed either. Since there is no difference of the behaviour of the SIP-AS when dealing with the service request for the Registered or Unregistered subscriber, it could be assumed that the mechanism of the SIP-AS restoration will be the same for the Registered or Unregistered subscriber.

The mentioned service request may be a third party registration request, a session origination service request or a session terminating service request. When there is a service request coming to the SIP-AS, the SIP-AS may be in failure status or has restored. If the SIP-AS has restored from the failure, the SIP-AS could obtain the subscriber service data from the HSS or other networks entities. For example, the restored SIP-AS receives a service request for a subscriber from the S-CSCF, but the SIP-AS has not the service data of this subscriber, then the SIP-AS could request the transparent data via Sh interface from the HSS in order to provide the service to the subscriber. So, the discussion on the failure of the SIP-AS is focused on the scenario that the SIP-AS is in the failure status.

When there is a service request coming to the SIP-AS, which is in failure, the following clauses cover the outcome in several different cases.

5.5.2 Third Party Registration Request

During the registration of the subscriber, the third party registration request will be triggered, and the S-CSCF will check whether the Filter Criteria of the subscriber matches the register request from the UE. If the SIP-AS matches the Filter Criteria of the subscriber service profile for the event of REGISTER request, the S-CSCF will initiate the third-party REGISTER request to the SIP-AS. If the SIP-AS were in failure, the SIP-AS could not respond the third party registration request, or the SIP-AS would return a failure response to the S-CSCF. According to the current 3GPP specification, the S-CSCF may abort sending the third-party REGISTER request, or initiate network-initiated deregistration procedure based on the information in the Filter Criteria. Although there may be more than one SIP-AS providing the same service, the subscriber can not register to the SIP-AS, and the related service can not be provided to the subscriber.

5.5.3 Originating Service Request

When the S-CSCF receives a session origination service request from the UE, the S-CSCF will check whether the Filter Criteria of the subscriber matches the service request. If the SIP-AS matches the Filter Criteria of the subscriber service profile for the event of service request, then the S-CSCF will forward the service request to the SIP-AS. If the SIP-AS were in failure, the SIP-AS could not respond the service request, or the SIP-AS would return a failure response to the S-CSCF. According to the current 3GPP specification, if the S-CSCF fails to receive a SIP response or receives a 408 (Request Timeout) response or a 5xx (Failure Response) response from the SIP-AS, the S-CSCF shall continue or terminate the session based on the default handling defined in the matched filter criteria. Although there may be more than one SIP-AS providing the same service, the subscriber can not register to the SIP-AS, and the related service can not be provided to the subscriber.

5.5.4 Terminating Service Request

When the S-CSCF receives a session terminating service request for the subscriber from the other end point, the S-CSCF will check whether the Filter Criteria of the subscriber matches the service request. If the SIP-AS matches the Filter Criteria of the subscriber service profile for the event of service request, then the S-CSCF will forward the service request to the SIP-AS. If the SIP-AS were in failure, the SIP-AS could not respond the service request, or the SIP-AS would return a failure response to the S-CSCF. According to the current 3GPP specification, if the S-CSCF fails to receive a SIP response or receives a 408 (Request Timeout) response or a 5xx (Failure Response) response from the SIP-AS, the S-CSCF shall continue or terminate the session based on the default handling defined in the matched filter

criteria. Although there may be more than one SIP-AS providing the same service, the subscriber can not register to the SIP-AS, and the related service can not be provided to the subscriber.

5.6 IP-CAN Service Interruption

Lack of IP connectivity will mean also lack of IMS service for the user. This situation will persist until the IP connectivity is restored. If the IMS-UE is able to retain its previous IP address, then it might be possible to resume service normally, otherwise a new registration will be required. If the problem affects a large number of users (e.g. SGSN or GGSN restart), there could be problems in the network due to the massive number of simultaneous registrations and IP changes from one user to another. Depending on the case, the user may or may not be aware of the problem unless it initiates a process that requires IP connectivity, which will fail. The procedures describing how to handle loss of IP connectivity to the terminal are in clauses 5.10.3 and E.2.1a.2 of 3GPP TS 23.228 [4] and in clause 5.2.8.1 of 3GPP TS 24.229 [5].

5.7 HSS Service Interruption

5.7.1 Introduction

This clause will analyse the impacts of a HSS stop in the network and in the service to the user in order to highlight the problems that need to be covered by the alternative solutions in this technical report. The initial state that will be considered is an IMS Core Network working properly and with ongoing traffic (a certain amount of users are registered and unregistered in the HSS). At one point the HSS stops operation, this implies lack of response from that HSS and potential loss of the temporary information of some subscribers in that HSS. Since only one HSS holds the subscriber data for that user, any service that requires intervention of the HSS will fail until that HSS resumes operation. The following clauses analyse the behaviour of the network once that the HSS is again in service (although some users' temporary data may be lost). The assumption is that registration state and S-CSCF name are temporary data as described in 3GPP TS 23.008 [6].
This scenario can be avoided if the HSS implements a regimen that secures both the permanent and the temporary data. This regimen can include replication of volatile storage units and periodic back-up of data to non-volatile storage. If the data security regimen ensures the integrity of the data in spite of failure of part of the HSS equipment then there will be no impact on service.

5.7.2 Unregistered User

Loss of registration state and S-CSCF name for an unregistered user may result in assignment of a different S-CSCF. There seems to be no difference in the service to the user for this reason, however as indicated in section 5.2.2 there could be problems in the SIP subscriptions to notifications (reg-event package) towards the previously assigned S-CSCF. In addition to that, the user will occupy memory space in the previously assigned S-CSCF until a cleanup takes place. The situation for originating unregistered requests is almost identical.

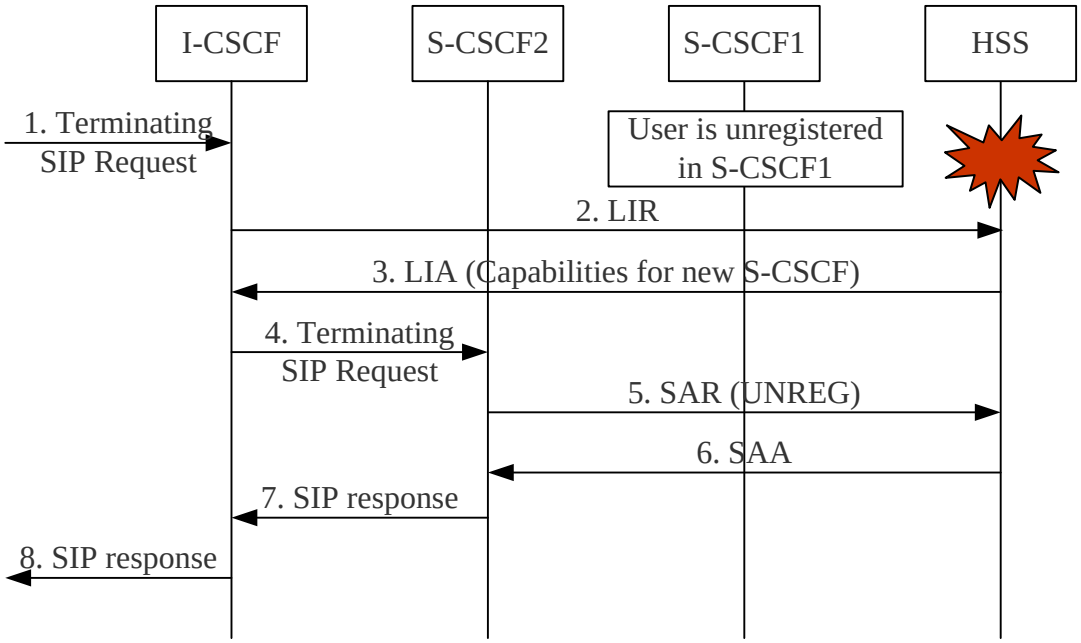

Figure 5.7.2.1. Terminating request to unregistered user with loss of HSS data

5.7.3 Originating Request from Registered User

Originating requests will be processed normally in the S-CSCF, even though the user data is lost in the HSS. Even a re-registration will be processed correctly unless a different S-CSCF has been assigned as described in clauses 5.7.2 and 5.7.4 and the re-registration is not authenticated. This last case of an authenticated re-registration will also be processed correctly (with a new S-CSCF assignment) because the S-CSCF name in the HSS will be overwritten with the MAR request.

5.7.4 Terminating Request to Registered User

Loss of registration state and S-CSCF name for a registered user may also result in assignment of a different S-CSCF when a terminating request is processed. In this case, the SIP request will be handled as for an unregistered user (assuming there are services associated with the unregistered state). The user will not be aware of this situation and will have no reason to think that those SIP requests are not being forwarded to the correct UE. This situation will remain until a re-registration is initiated from the UE. Considering that the re-registration timers can be quite long the consequence of this can be seen as quite severe and there is a need to try to improve the service availability in this scenario.

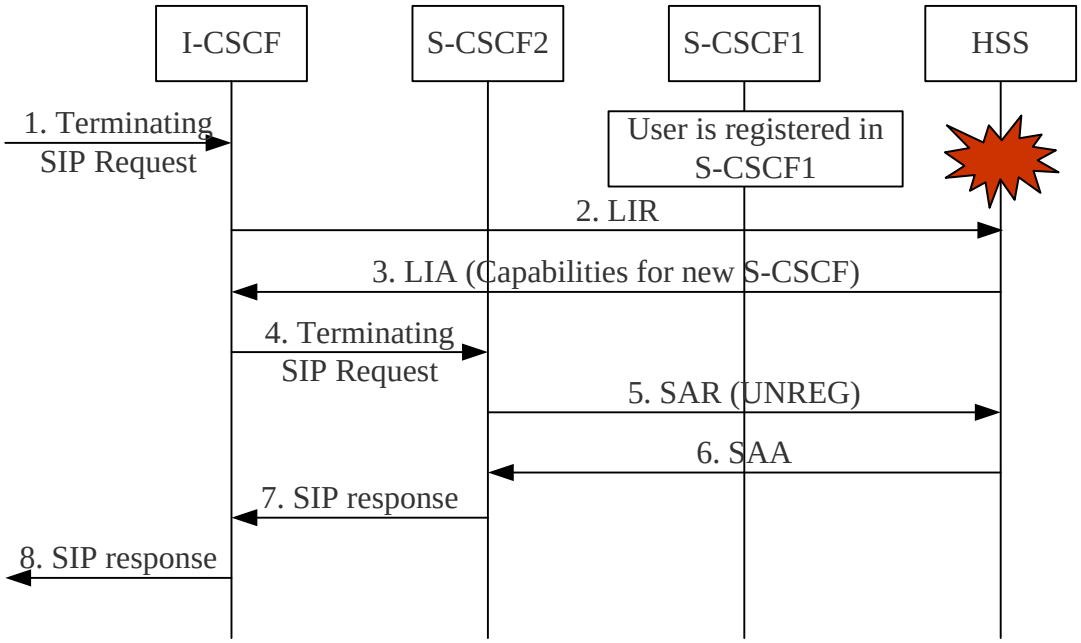

Figure 5.7.4.1. Terminating request to registered user with loss of HSS data

5.7.5 Impacts on the Sh Interface

The HSS may also share data with the SIP-AS, and these data could be affected by an interruption of service of the HSS. This refers to the transparent data that the SIP-AS may have stored in the HSS and the subscriptions to notifications using the Sh Interface. If the subscriptions to notifications are lost, the SIP-AS will not receive the notifications that it expects until it subscribes again. The problem is even more serious for transparent data, which might be lost permanently.

6 Alternative solutions

6.1 Backup of S-CSCF Information in the HSS

6.1.1 Introduction

This solution for processing of SIP requests after a S-CSCF restart or during a S-CSCF downtime is based in the principle that S-CSCF information is stored in the HSS. There is a set of basic information that would be required for an S-CSCF to handle the requests for a user; it comprises the list of SIP Proxies (usually the P-CSCF address), list of Public User Identities, Contact Addresses and Contact Header parameters. In order to request the user to authenticate, the information of the UE's subscription to the reg-event package may also be required to be stored in the HSS. Other information, such as active subscriptions using the SIP reg-event package could be refreshed by the clients of the subscription after take over. The following clauses describe the modifications required for this solution.

6.1.2 Normal Registration

In order for the information to be available later, the S-CSCF would be required to store it in the HSS. This could be done during the registration process with an additional information element in the SAR request. In addition to the basic set of information required to handle traffic. The changes to the protocol would be in the form of an additional information element in table 6.1.2.1 of 3GPP TS 29.228 [10]:

S-CSCF Information (See 7.X)	SCSCF-Information	O	If Server-Assignment-Type is REGISTRATION or RE_REGISTRATION, the S-CSCF may optionally send this information element to the HSS. This information allows a later retrieval in case of an S-CSCF service interruption.

The detailed behavior in clause 6.1.2.1 should also indicate that the HSS shall store this information if it implements IMS Restoration Procedures. The information will be associated to the Private User Identity and the Implicit Registration Set that is affected by the SAR request, and it will contain:

- the list of SIP proxies in the path (normally it would be just the P-CSCF address),
- the Contact Information (Contact Addresses and Contact Header parameters),

NOTE: If the solution included the need to send a NOTIFY to the UE, then the UE's subscription information (Call-ID, From, To, Record-Route) would also need to be stored. To avoid frequent storing of the subscription information in the HSS, the CSeq should not be included in the S-CSCF Information.

As the following figure 6.1.2.1 shows, the S-CSCF backups user registration related data such as the content of Path, Contact address in the HSS during the initial registration procedure. When the S-CSCF sends SAR to request the user profile from the HSS in step 14 after successful authentication, the backup data are carried in the message to the HSS at the same time.

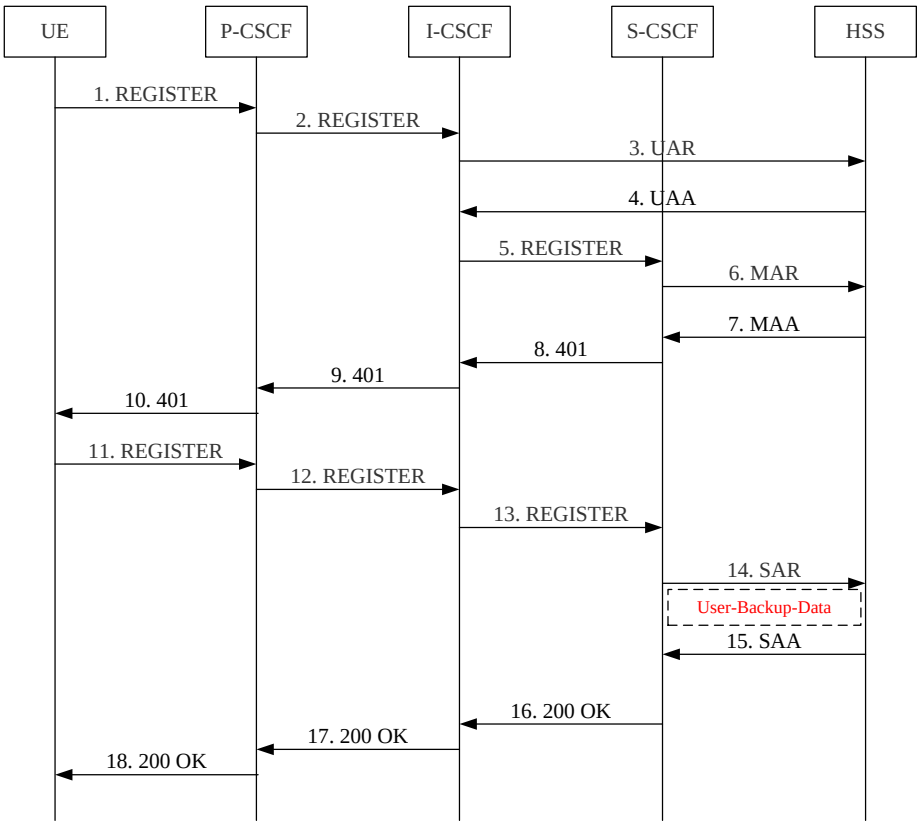

Figure 6.1.2.1: Backup of registration data during initial registration procedure

1-13. The UE initiates the registration towards the network and the normal registration procedure follows.
14. The S-CSCF name is changed and the S-CSCF sends the SAR with these registration related data.
15. The HSS stores the registration related data and returns the SAA to the S-CSCF with service profile.
16-18. Normal registration procedure.

Note 1: In the above step 14, it is needed to extend the function of the HSS and the S-CSCF and enhance the Cx interface to make it possible that the S-CSCF could backup registration data such as Path and Contact address of the user in the HSS during the registration procedure.

Note 2: It will be possible for the S-CSCF to update the S-CSCF Information in the HSS at any time using an SAR with Server Assignment Type RE_REGISTRATION or a new Server Assignment Type specific for this purpose. This could be used, for example to update the value of Temporary GRUUs assigned in the S-CSCF.

6.1.3 Retrieval of S-CSCF Information

According to the analysis in clause 5.2, the most relevant case in which the S-CSCF might need this information is when processing a terminating SIP Request. In this case, when the HSS believes the identity to be Registered, the SAA response will be either DIAMETER_ERROR_IDENTITY_ALREADY_REGISTERED if the request is originated in a different S-CSCF or DIAMETER_ERROR_IN_ASSIGNMENT_TYPE if it comes from the same S-CSCF.
For DIAMETER_ERROR_IN_ASSIGNMENT_TYPE, the proposal is to add the information stored in the HSS to the response so that the S-CSCF has the possibility to forward the request to the P-CSCF that attends the user. The other error code covers the case of a double S-CSCF assignment. To avoid that situation, the proposal is to create a new value

for Server-Assignment-Type in order to overwrite the S-CSCF name stored in the HSS and to allow this operation only if there was a previous request for capabilities from the I-CSCF. As an alternative to creating a new value for Server-Assignment-Type the existing value (UNREGISTERED) could be used when the HSS is able to recognize whether or not the stored S-CSCF is contactable. The I-CSCF should also be allowed to request REGISTRATION_AND_CAPABILITIES in the terminating case. The changes to the protocol would be in table 6.1.2.2 adding the information element with S-CSCF information:

S-CSCF Information (See 7.X)	SCSCF-Information	C	The HSS shall send this information element if it implements the IMS Restoration Procedures and Experimental-Result is DIAMETER_ERROR_IN_ASSIGNMENT_TYPE or the value of Server-Assignment-Type in the request is NO_ASSIGNMENT or REASSIGNMENT.

The following figure shows the traffic flow for retrieval of information from the same S-CSCF:

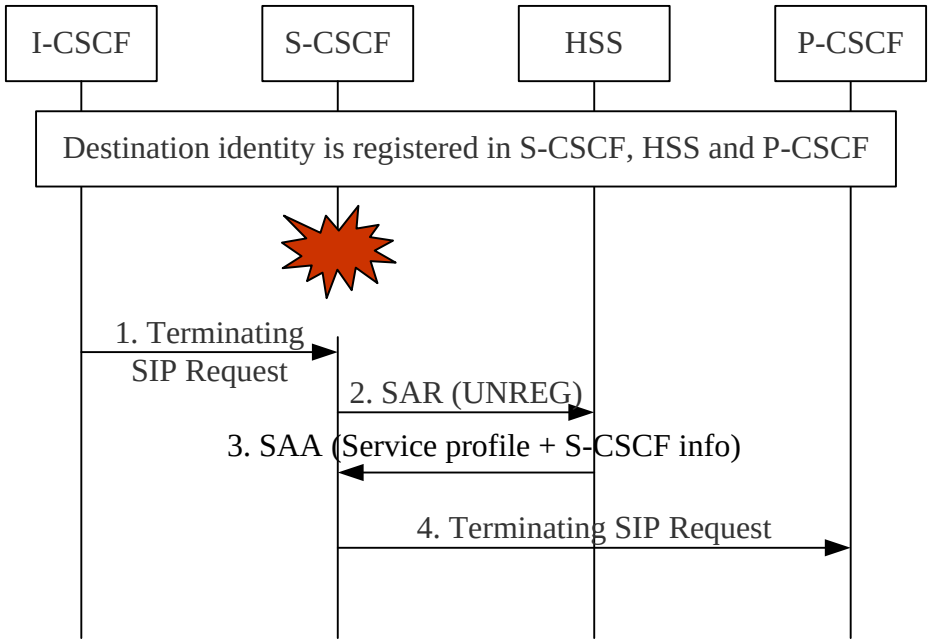

Figure 6.1.3.1: Retrieval from the previous S-CSCF

Note 1: In the above steps 2 and 3, it is needed to send the S-CSCF information from the HSS to the S-CSCF. This means an enhancement of both HSS and S-CSCF and it can be done in two ways. One is enhancing the S-CSCF, so that when it receives the SAA with Experimental-Result-Code value set to DIAMETER_ERROR_IN_ASSIGNMENT_TYPE, it is able to send SAR with a new Server Assignment Type as REASSIGNMENT to request backuped data from the HSS. The other is represented in figure 6.1.3.1 and it is enhancing the HSS to add the information in the SAA for this error code.

Note 2: The decision of when to assign a new S-CSCF for the user is taken in the I-CSCF and how to make that decision is left as an implementation option.

Note 3: In the above step 3, if the related Public User Identity is registered and the HSS does not have the S-CSCF info, the HSS may download the service profile in the SAA with Experimental-Result-Code value set to DIAMETER_ERROR_IN_ASSIGNMENT_TYPE.

Note 4: In the above step 3, if the S-CSCF only receives service profile in the SAA, i.e. no S-CSCF info is carried in the SAA, the S-CSCF shall only trigger matched unregistered service for the user. Otherwise, if the S-CSCF receives service profile together with the S-CSCF info, the S-CSCF shall trigger matched registered service for the user.

The procedures shown in the Figure 6.1.3.1 could also be applied to the originating request initiated by a SIP-AS on behalf of an unregistered or registered user with terminating SIP request replaced by Originating SIP request and P-CSCF replaced by terminating network. There is another little difference in the S-CSCF about triggering registered or unregistered service. If the S-CSCF receives service profile in the SAA with Experimental-Result-Code value set to DIAMETER_ERROR_IN_ASSIGNMENT_TYPE, the S-CSCF shall trigger matched registered service for the user. Otherwise, if the S-CSCF receives service profile in the SAA with Experimental-Result-Code value set to DIAMETER_SUCCESS, the S-CSCF shall only trigger matched unregistered service for the user as the current procedure.

The following figure shows the traffic flow for retrieval of information from a different S-CSCF:

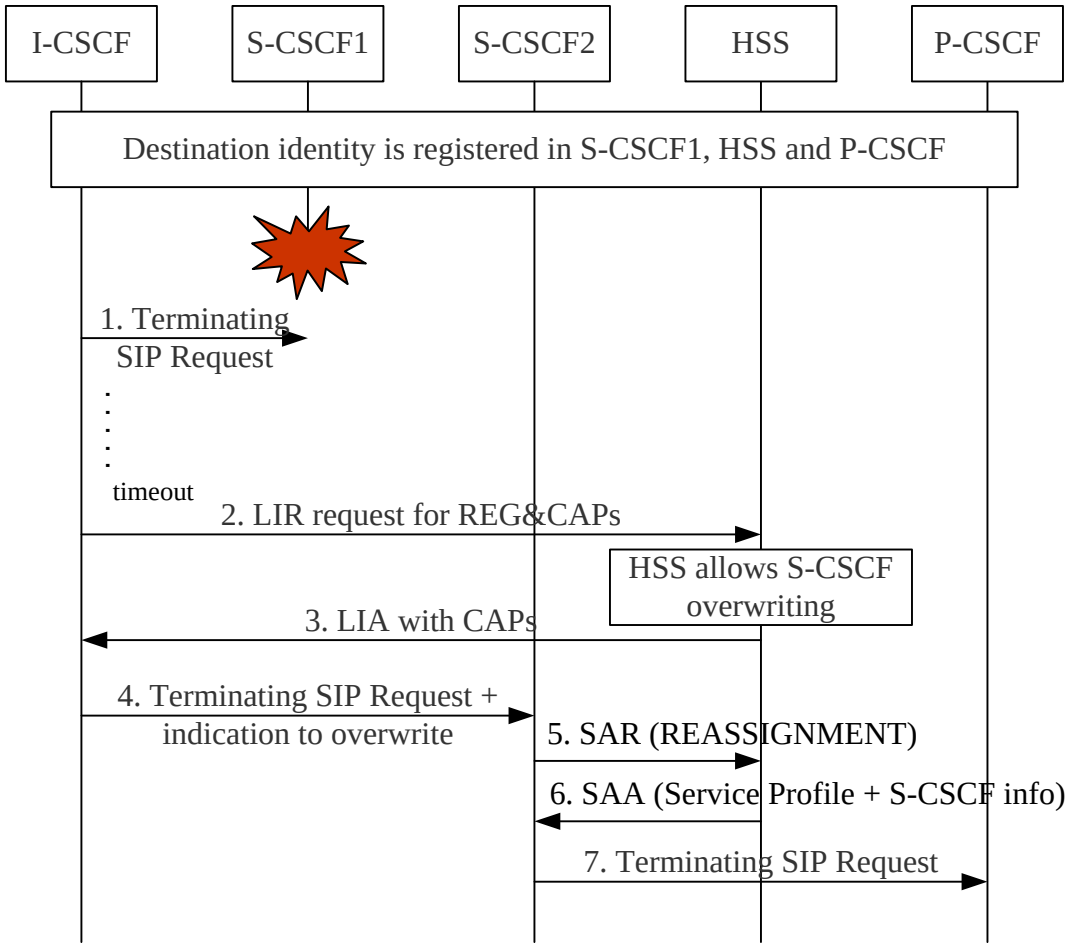

Figure 6.1.3.2: Retrieval from a different S-CSCF

NOTE 1: The failure of the S-CSCF1 may be detected before the I-CSCF forwards the INVITE to the S-CSCF1.
NOTE 2: In the above step 2, it is needed to enhance the I-CSCF to be able to explicitly request S-CSCF capabilities and re-select the S-CSCF during terminating procedure.
NOTE 3: In the above step 4, the I-CSCF may carry an indication for restoration in the terminating SIP request to indicate the S-CSCF it is a new assigned server caused by the failure of the originally assigned S-CSCF. The way in which this indication might be put into the SIP signalling is out of the scope of this document. Adding a restoration indication to the SIP request is feasible but requires work within IETF.
NOTE 4: In the above step 5, it is needed to extend the SAR with a new Server Assignment Type as REASSIGNMENT to request backuped data from the HSS.
NOTE 5: In the above step 6, it is needed to enhance the HSS to be able to send the user profile even if the requesting S-CSCF is different from the stored S-CSCF.
NOTE 6: In the above step 6, it is needed to extend the SAA and enhance the HSS to be able to send the backuped data together with the user profile after receiving REASSIGNMENT indication from the S-CSCF, even if the requesting S-CSCF is different from the stored S-CSCF.
NOTE 7: In the above step 6, if the related Public User Identity is registered and the HSS does not have the S-CSCF info, the HSS may download the service profile in the SAA.
NOTE 8: In the above step 6, if the S-CSCF only receives service profile in the SAA, i.e. no S-CSCF info is carried in the SAA, the S-CSCF shall only trigger matched unregistered service for the user. Otherwise, if the S-CSCF receives service profile together with the S-CSCF info, the S-CSCF shall trigger matched registered service for the user.

The procedures shown in the Figure 6.1.3.1 could also be applied to the originating request initiated by a SIP-AS on behalf of an unregistered or registered user with terminating SIP request replaced by Originating SIP request and P-CSCF replaced by terminating network.

Another alternative would be to trigger an initial registration of the UE after step 3 (when the request using the retrieved S-CSCF information is finished), so that the status of the registered user can be rebuilt completely across the network. This can be done with a NOTIFY using the UE subscription information to the reg-event package that was previously stored in the HSS.

NOTE: Since the retrieved subscription information does not include the CSeq parameter, the S-CSCF could use an allowed maximum CSeq value in the NOTIFY message.

The following figure shows an alternative solution traffic flow for retrieval of information from a different S-CSCF:

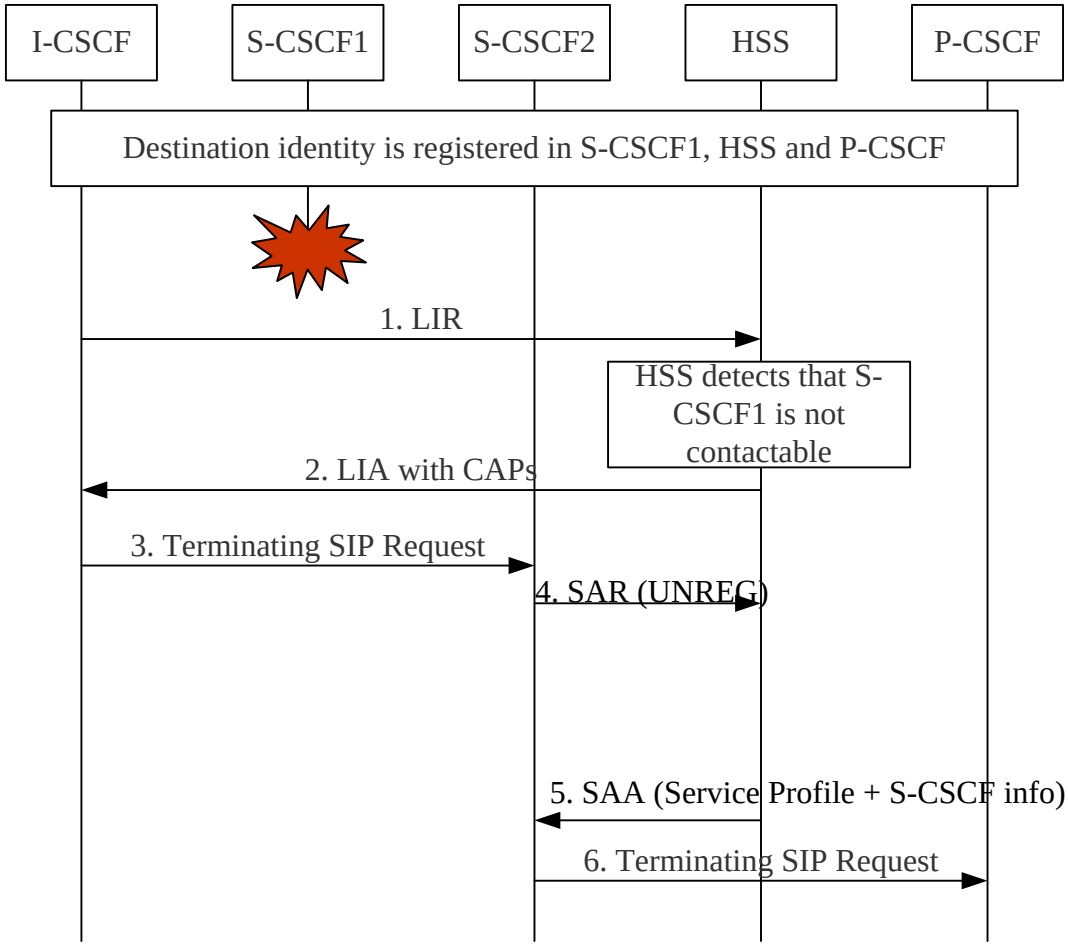

Figure 6.1.3.3: Alternative Retrieval from a different S-CSCF

NOTE 1: The failure of the S-CSCF1 may be detected by the HSS by implementation specific means which may e.g. be based on existing Diameter mechanisms.

NOTE 2: In the above step 2, it is needed to enhance the HSS to return capabilities although name of S-CSCF1 is stored and when the stored S-CSCF is not contactable.

NOTE 3: In the above step 4, it is needed to enhance the HSS to accept the server assignment type UNREG since the stored S-CSCF1 is not contactable, and to return Service Profile + S-CSCF info to S-CSCF2.

6.1.4 Removal of S-CSCF Information

The S-CSCF information that is stored in the HSS during the registration needs to be deleted when the identity is no longer in Registered State. Deregistration may be user-initiated or network-initiated. The following figures shows the traffic flow for removal of the S-CSCF Information in the two deregistration cases:

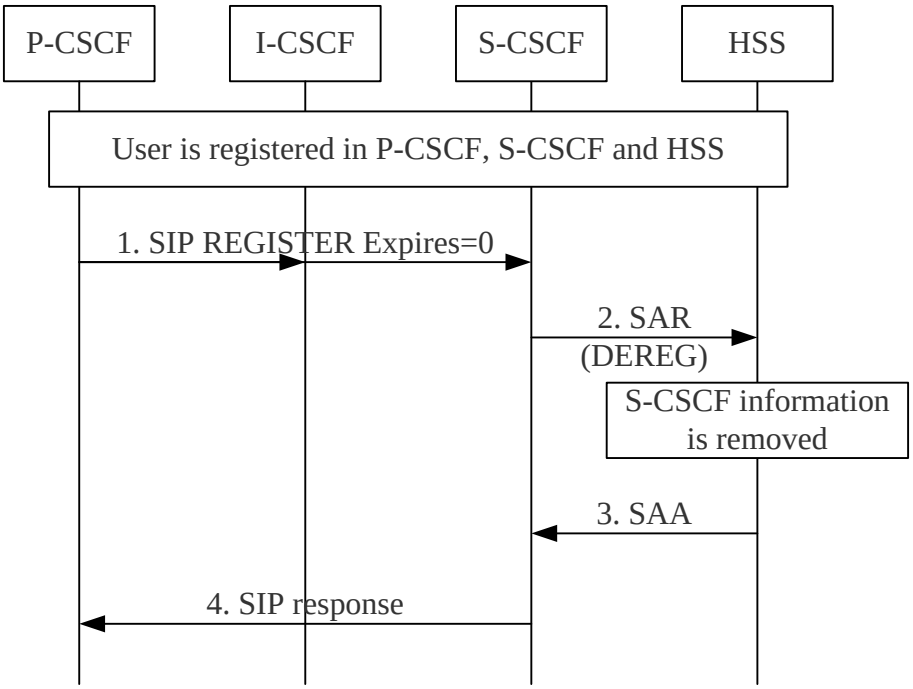

Figure 6.1.4.1: Removal of S-CSCF information during user-initiated deregistration

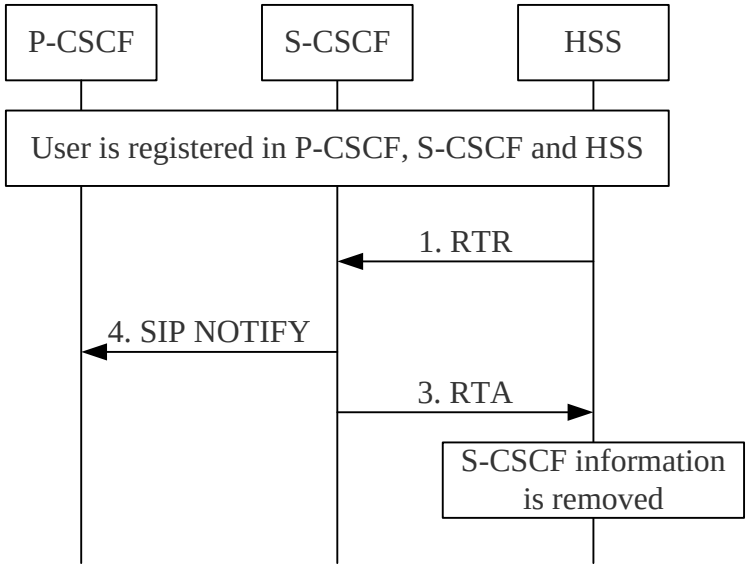

Figure 6.1.4.2: Removal of S-CSCF information during network-initiated deregistration

6.1.5 Handling of Originating SIP Request for an Unknown User in the S-CSCF

Since we are working with the assumption that the S-CSCF could lose the information related with some identities, it may also happen that the S-CSCF receives an originating request for a user that is not known in that S-CSCF. If that is the case, the S-CSCF should check if there is any information related to that user in the HSS. The proposal is to perform this checking using the NO_ASSIGNMENT Server-Assignment-Type, which will also return the S-CSCF information. The following figure shows the traffic flow for retrieval of information when processing an originating SIP Request:

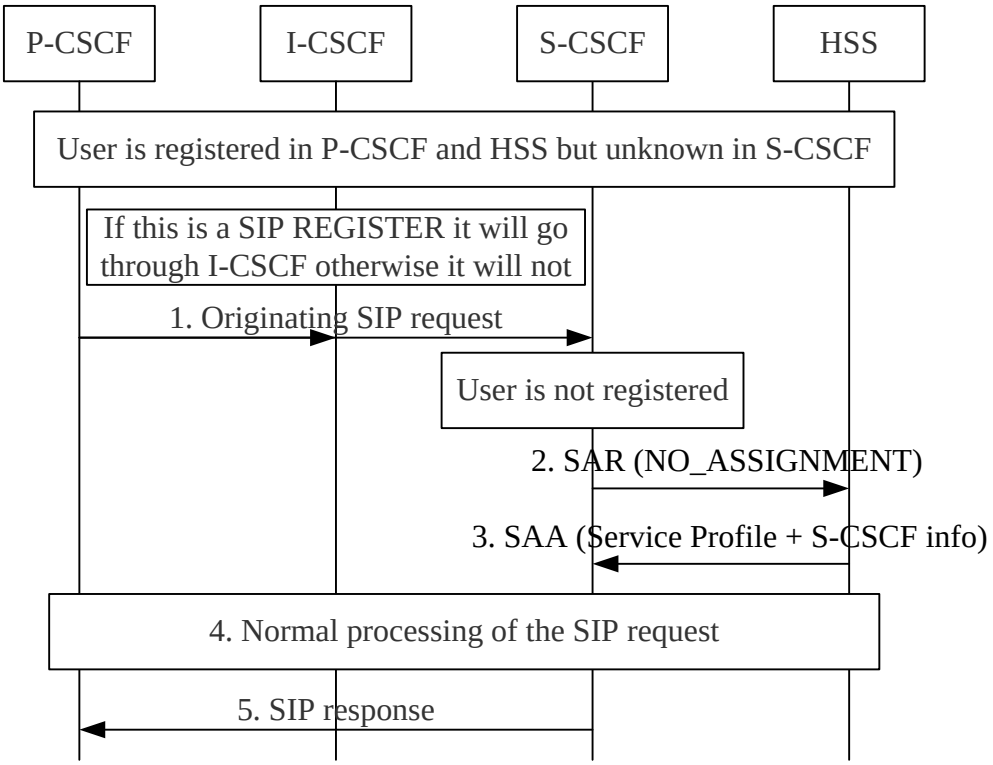

Figure 6.1.5.1: Retrieval of information for an unknown user during originating SIP request

NOTE 1: In the above steps 2 and 3 the S-CSCF information is not needed to process the originating SIP request, but it could be used by the S-CSCF in subsequent terminating SIP requests.

NOTE 2: The above step 3 assumes that S-CSCF information is always sent in the responses to an SAR with Server Assignment Type NO_ASSIGNMENT. The decision on whether it would be more appropriate to have a new Server Assignment Type to request this information is out of the scope of this document.

6.1.6 Restoration During Re-registration Procedure

When the I-CSCF receives a re-registration session, but the assigned S-CSCF retrieved from the HSS can not be contacted, one possible restoration solution is as the following.

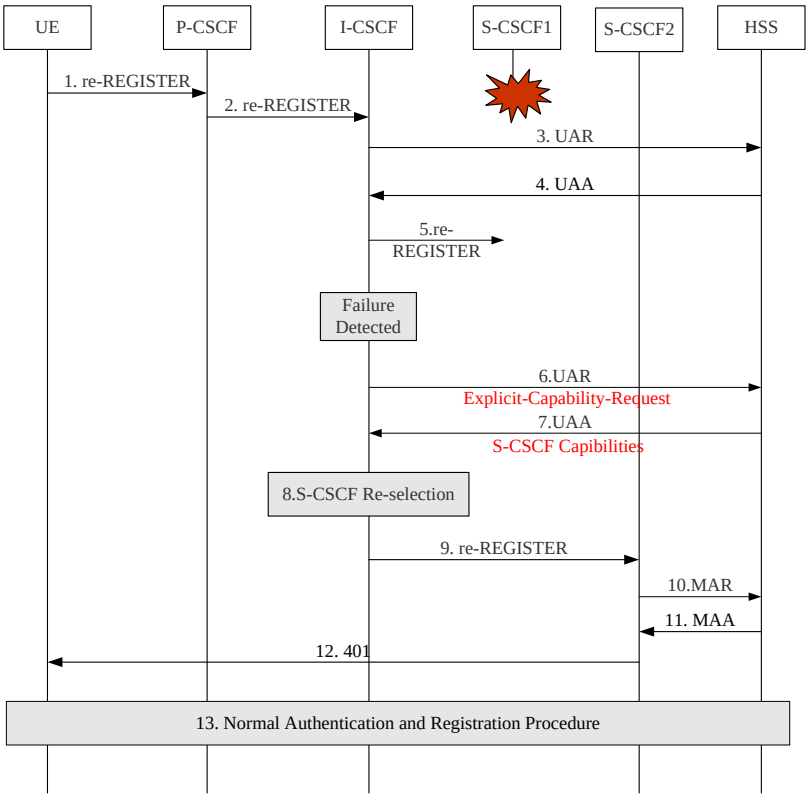

Figure 6.2.2.3.1 Restoration during re-registration procedure

1. The UE sends a re-REGISTER request to the P-CSCF.
2. The P-CSCF forwards the re-REGISTER request to the I-CSCF.
3. The I-CSCF sends UAR to the HSS to request the S-CSCF name.
4. The HSS sends back the UAA with the S-CSCF1 name.
5. The I-CSCF forwards the re-REGISTER to the S-CSCF1, but does not receive any response from the S-CSCF1.

NOTE 1: The failure of the S-CSCF1 may be detected before the I-CSCF forwards the re-REGISTER to the S-CSCF1.

6. The I-CSCF sends UAR to the HSS to explicitly queries the S-CSCF capabilities.
7. The HSS sends back the UAA with the S-CSCF capabilities.
8. The I-CSCF re-selects a new S-CSCF2 for the user.
9. The I-CSCF forwards the re-REGISTER to the S-CSCF2.
10. The S-CSCF2 sends MAR to request the user's authentication data.
11. The HSS sends back the MAA with the authentication data.

NOTE 2: Above steps 10 and 11 will be skipped if the re-registrations are not authenticated.

12. The S-CSCF2 sends 401 towards the UE.
13. Normal authentication and registration procedure.

NOTE 3: In the above step 6, it is needed to extend the function of the I-CSCF to be able to explicitly request S-CSCF capabilities and re-select the S-CSCF during re-Registration procedure, just as it does during the initial registration procedure. In other words, the I-CSCF could re-select the S-CSCF without checking whether this is an initial registration or re-registration.

NOTE 4: The failure scenario corresponding to the procedure above will be handled with the current procedures with the return of an error, requiring a new initial registration by the user, which also restores the network to a consistent state. This change avoids the need to return that error and the subsequent initial registration.

NOTE 5: In the above step 13, it is needed to extend the function of the HSS to enable overwriting of the S-CSCF name with an SAR after an explicit request from the I-CSCF of the capabilities for that user.

6.1.7 Backup of S-CSCF information after UE's subscription

When the S-CSCF receives the UE's subscription to notification of the reg-event, the S-CSCF could send an SAR carrying the information of the UE's subscription to the HSS for backup. The value of the Server Assignment Type could be set to RE_REGISTRATION. The User Data Already Available parameter should be set to USER_DATA_ALREADY_AVAILABLE.

NOTE: The S-CSCF would always send an SAR to the HSS to store the UE's subscription information when this is the first subscription of the UE in the S-CSCF. In other cases, the S-CSCF could compare the subscription information with the existing one. If the subscription information including Call-ID, From, To, Record-Route does not change, the S-CSCF need not send SAR for backup.

The S-CSCF should send the registration data together with the subscription data as one S-CSCF information, in order to avoid additional handling in the HSS. Each time the HSS receives the S-CSCF information related to the same Private User Identity in the SAR, the HSS just overwrite the old one. The procedure is showed in the following figure.

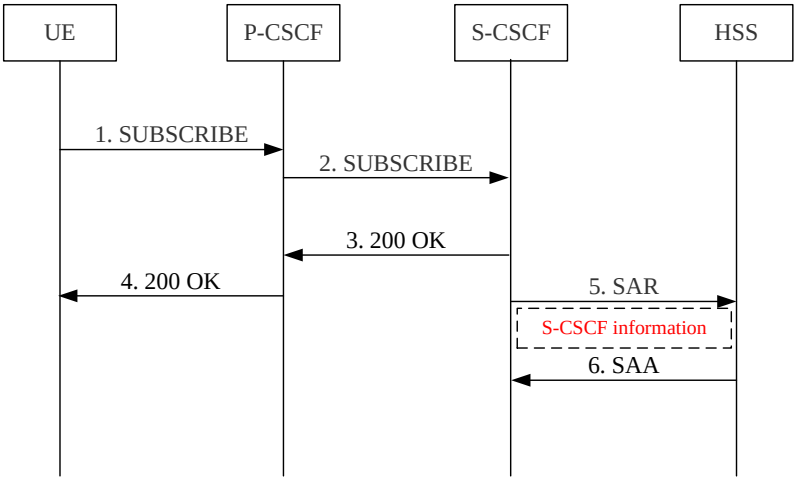

Figure 6.1.7.1 Backup data after UE's subscription

NOTE: The S-CSCF information in the step 5 shall include registration data and subscription data.

6.2 Triggering of initial registration from the S-CSCF or P-CSCF

6.2.1 Introduction

This procedure addresses the processing of originating SIP requests after a S-CSCF restart or during a S-CSCF downtime. It is based on the use of a specific SIP error code that shall trigger the initial registration of the user.

NOTE: It could be seen that currently the solutions will only take effect for Rel-8 IMS UEs.

6.2.2 Originating Traffic Restoration

6.2.2.1 When the Assigned S-CSCF Is Unavailable

When the P-CSCF receives an originating session request, but the S-CSCF in the Route can not be contacted, one possible restoration solution is as the following.

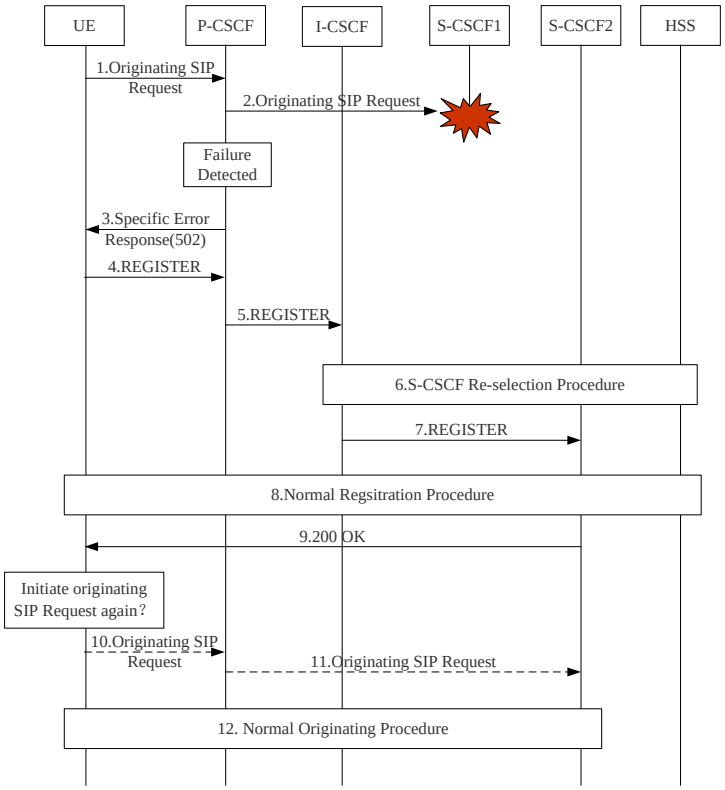

Figure 6.2.2.1.1 Restoration during originating procedure

1. The UE sends an originating SIP Request to the P-CSCF.
2. The P-CSCF forwards the originating SIP Request to the S-CSCF1 according to the Route header, but does not receive any response from the S-CSCF1.

   NOTE 1: The failure of the S-CSCF1 may be detected before the P-CSCF forwards the Originating SIP Request to the S-CSCF1.

   NOTE 2: How the P-CSCF takes the decision to consider that the S-CSCF has failed and to return the error is left as an implementation option.
3. The P-CSCF returns a specific error response such as 502 (Bad Gateway) to the UE.
4. The UE receives the specific error response, and it shall initiate a new registration towards the network immediately.
5. The P-CSCF forwards the message to the I-CSCF as normal.
6. The I-CSCF interacts with the HSS and re-selects a new S-CSCF2 for the user.

7-9. The I-CSCF forwards the message to the S-CSCF2 and the normal registration procedure follows.

10. After receiving the 200 OK for the registration request, the UE may decide to send the originating SIP Request again.
11. The P-CSCF forwards the Originating SIP Request to the S-CSCF2.
12. Normal originating procedure.

    NOTE 3: In the above step 3 and step 4, the P-CSCF returns a specific error response indicating that the UE shall initiate a registration towards the network, which may need to extend the function of the P-CSCF and the UE.

6.2.2.2 Restoration When the Assigned S-CSCF Resumes

When the P-CSCF receives an originating session request, but the S-CSCF in the Route restarted before and has not the user data or the data is not available, one possible restoration solution is as the following.

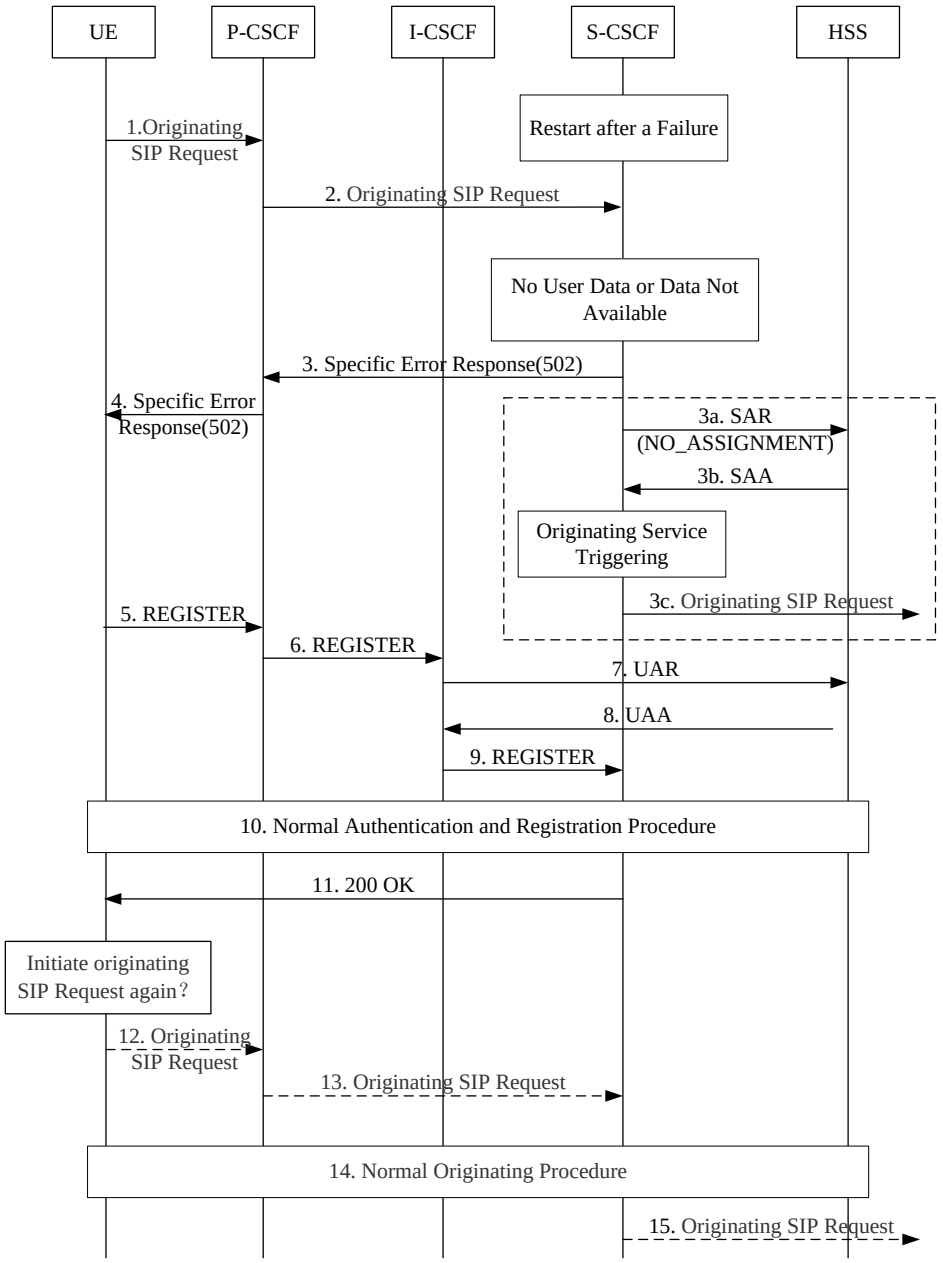

Figure 6.2.2.2.1 Restoration during originating procedure

1. The UE initiates an originating SIP Request request.
2. The P-CSCF forwards the originating SIP Request to the S-CSCF according to the Route header.
3. The S-CSCF does not have the user's data or the user data is not available, returns a specific error response such as 502 (Bad Gateway) to the P-CSCF.

3a-3c. Another way to do this is according to 6.1.5.

4. The P-CSCF forwards the error response to the UE.
5. The UE receives the error response, and it shall initiate a new registration towards the network immediately.

6-11. Normal authentication and registration procedure.

12. After receiving the 200 OK for the registration request, the UE may decide to initiate the originating SIP Request again.
13. The P-CSCF forwards the originating SIP Request to the S-CSCF.
14. Normal originating procedure.
15. The S-CSCF sends the originating SIP Request towards the terminating network.

NOTE: In the above step 3, 4 and 5, the S-CSCF returns a specific error response indicating that the UE shall initiate a registration towards the network, which may need to extend the function of the S-CSCF and the UE.

6.2A Precautionary de-registration of un-registered users

To avoid loss of service for un-registered users (see chapter 5.2.2) the S-CSCF may precautionary de-register the un-registered user when termination of all ongoing sessions is detected. The de-registration may be delayed by a configurable timer. Short timer values will increase de-registration traffic and registration traffic; on the other hand chances that the user suffers from loss of service will decrease.

6.2B S-CSCF re-assignment for unregistered user

To address the problem outlined in figure 5.2.2.1 a solution similar to that shown in figure 6.1.3.2 and 6.1.3.2.a is proposed i.e. a new S-CSCF is selected. The HSS needs to indicate to the new S-CSCF that the user is un-registered by not sending S-CSCF info (backup data) within SAA.

To address the problem outlined in figure 5.2.2.2 it is proposed that the SIP-AS after timeout resends the Originating SIP Request to an I-CSCF (which will then be responsible for S-CSCF re-selection).

To address the problem outlined in figure 5.2.2.3 the same solution as for the problem in figure 5.2.2.1 applies.

6.3 Second P-CSCF and deregistration from S-CSCF

One possible solution for the P-CSCF service interruption is that the P-CSCF adds a second P-CSCF into the Path along with its own address when sending the register message to the S-CSCF.

NOTE : How the second P-CSCF is selected is FFS. One possible way is that it could be pre-configured in the registered P-CSCF. The second P-CSCF can be found by the UE through the P-CSCF discovery procedure or it can be selected by the IMS network during the UE's initial registration procedure.

When the P-CSCF fails the S-CSCF is able to send Notify to the UE to initiate a new register through the second P-CSCF by the UE's subscription to the reg-event package. Since the second P-CSCF does not have the user data and security associations with the UE, it only needs to forward the Notify message to the UE without protection. When the UE receives the Notify not protected and even not from the P-CSCF it stored, but with the same subscription information such as CALL-ID it has, and because the subscription to the reg-event package is sent on the security association and no other entity could get this information, the UE could just trust it for this time and initiate a new registration. After the normal registration, the S-CSCF forwards the terminating call to the P-CSCF indicated in the Path. The restoration is shown as the following:

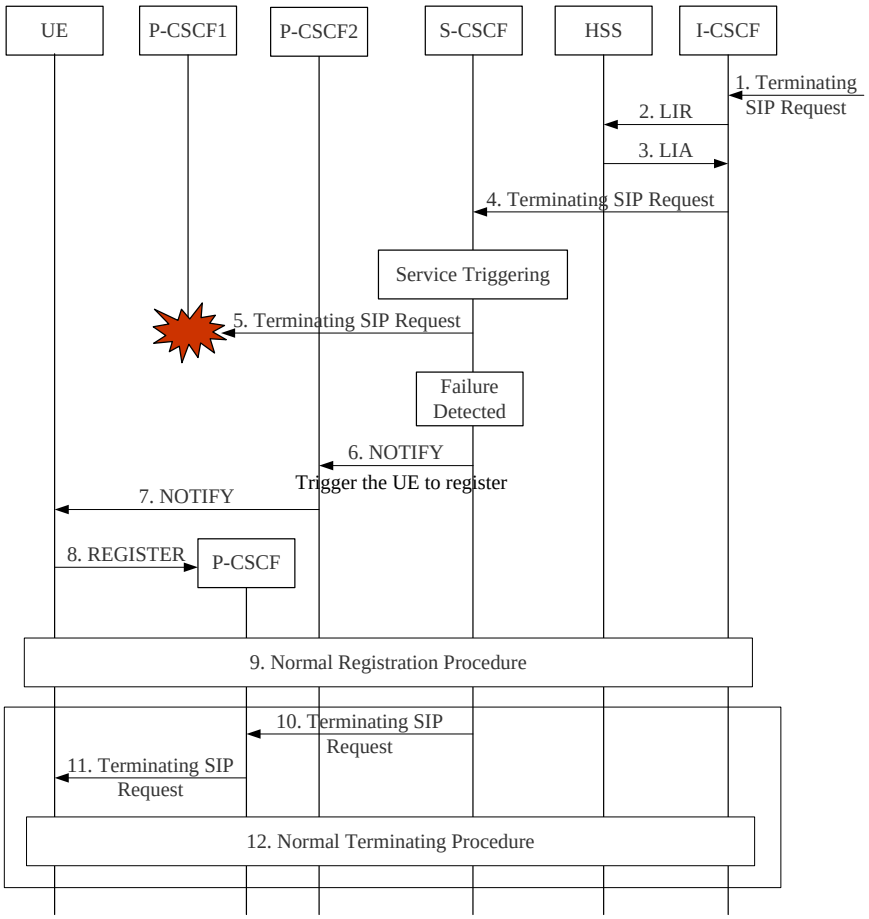

Figure 6.3.1 Restoration during terminating procedure

1. The I-CSCF receives a terminating SIP Request.
2. The I-CSCF sends LIR to the HSS to request the S-CSCF name.
3. The HSS sends back the LIA with the S-CSCF name.
4. The I-CSCF forwards the terminating SIP Request to the S-CSCF. The S-CSCF triggers the service for the user.
5. The S-CSCF forwards the terminating SIP Request to the P-CSCF1, but does not receive any response from the P-CSCF1.
6. The S-CSCF sends NOTIFY to the P-CSCF2 which was stored together with the P-CSCF1 received in the Path header during registration to trigger the UE to initiate a new register.
7. The P-CSCF2 forwards the NOTIFY to the UE without protection.
8. The UE checks the message, if the subscription information is the same as the UE stored, initiates a new registration.
   Note: The P-CSCF to which the UE sends the new REGISTER message may be the P-CSCF2 or the other P-CSCF which the UE finds during the normal P-CSCF discovery procedure.
9. Normal registration procedure.

10-12. The S-CSCF may forward the terminating SIP Request to the P-CSCF indicated in the Path, and continue the normal terminating procedure.

The above procedure could also be used with a little change for service restoration if the P-CSCF that receives terminating request from the S-CSCF does not have the user data. In such a case, the P-CSCF could return a specific error to the S-CSCF. Then the S-CSCF constructs the NOTIFY message and sends by the original P-CSCF or the second P-CSCF mentioned above to trigger the UE to initiate a new register. The enhancement needed to the P-CSCF and UE about forwarding and receiving NOTIFY without protection is the same as the above.

Note 2: Sending a NOTIFY that does not have information about the SUBSCRIBE that originated that NOTIFY was considered as a potential door to a Denial of Service attack by SA3. This could be avoided if the operators could ensure their underlying access network is secure, e.g., by enabling the GPRS or UMTS encryption, or by activating IMS confidentiality protection.

6.3.1 Select Second P-CSCFs for the usage of restoration

The UE can negotiate the second P-CSCFs for the usage of restoration with the IMS network. It can be done during the UE's initial registration procedure. The UE and the IMS network (eg. the S-CSCF) should store this second P-CSCF information. When the working P-CSCF fails, the S-CSCF can inform and force the UE to perform an initial registration through the previous selected second P-CSCF for restoration.
The second P-CSCF can be found by the UE through the P-CSCF discovery procedure or it can be selected by the IMS network during the UE's initial registration procedure.
If the UE gets the second P-CSCF for restoration, it may carry this special P-CSCF information to the IMS network during the initial registration. And the IMS network should store it.
If, during initial registration, the UE doesn't provide the second P-CSCF for restoration, the IMS network should selects someone for the UE, and notify the UE the selected P-CSCF. Both the S-CSCF and the UE should store this second P-CSCF information.
The initial registration procedure including second P-CSCF selected for restoration, is described as follows.

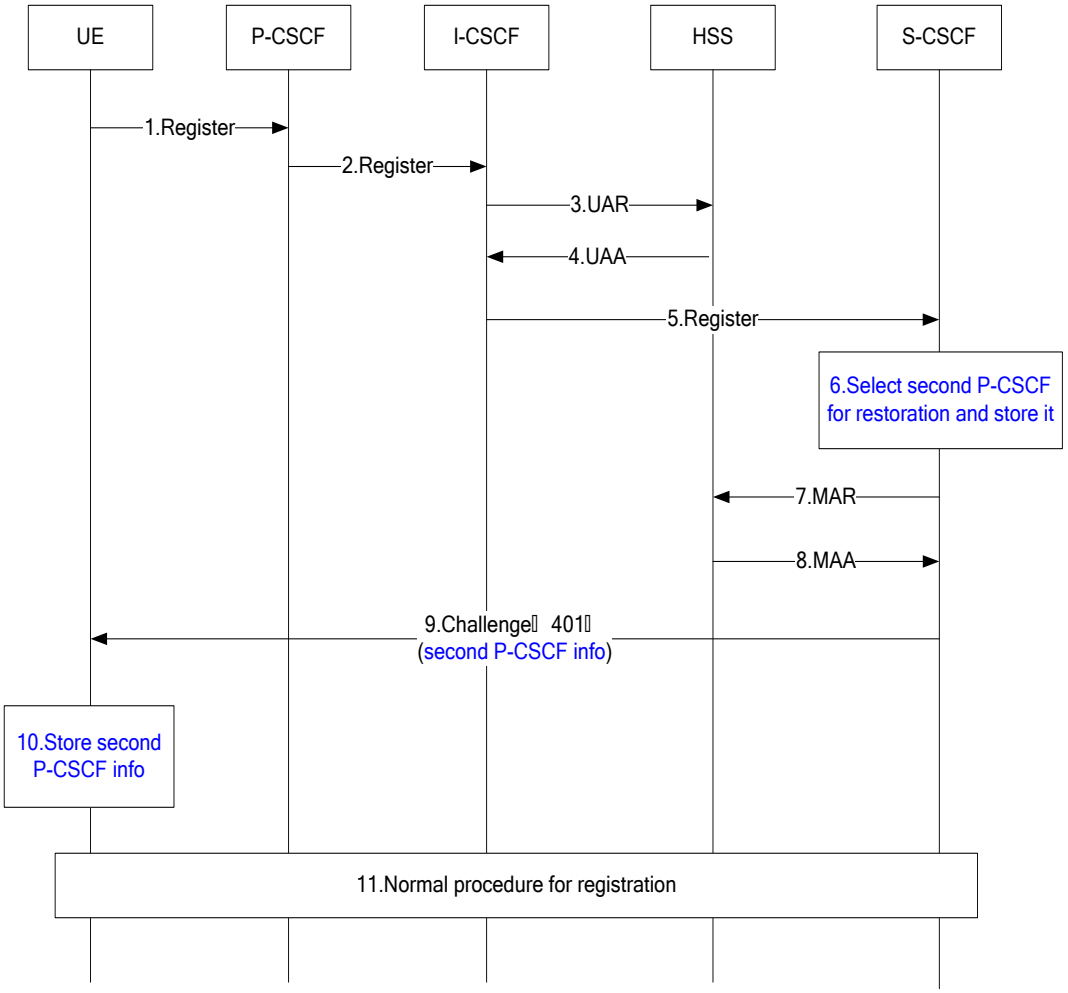

Figure 6.3.1.1 selecting second P-CSCF for restoration during initial registration procedure

1. The UE performs an initial registration, without carrying any additional P-CSCF information for restoration. The UE sends a REGISTER request to the P-CSCF.
2. The P-CSCF forwards the REGISTER request to the I-CSCF.
3. The I-CSCF sends UAR to the HSS to explicitly query the S-CSCF capabilities.
4. The HSS sends back the UAA with the S-CSCF capabilities.
5. The I-CSCF selects an S-CSCF for the user, and forwards the REGISTER to the S-CSCF.
6. The S-CSCF checks that there is no second P-CSCF for restoration carried in the SIP signalling, so the S-CSCF selects and stores some one P-CSCF for the UE as restoration usage.

NOTE 1: How to select the second P-CSCF is FFS. One possible way is that it could be pre-configured in the registered P-CSCF.

NOTE 2: Both the S-CSCF and the P-CSCF in the registration procedure can select the second P-CSCF for restoration. The working P-CSCF can select a second P-CSCF and add this information in the path along with itself, and forwarding this information to the S-CSCF. The working S-CSCF can select second P-CSCF and don't need to add this information to the signalling path.

NOTE 3: How network entities select the second P-CSCF for restoration can be considered as implementation work, since the working P-CSCF may select visited network P-CSCF as restoration one, and the S-CSCF may select home network P-CSCF.

7. The S-CSCF sends MAR to request the user's authentication data.
8. The HSS sends back the MAA with the authentication data.
9. The S-CSCF sends 401 towards the UE, including the selected second P-CSCF for restoration.
10. The UE retrieves and stores the second P-CSCF for restoration usage.
11. Normal authentication and registration procedure.

NOTE 4: How to inform the UE the selected second P-CSCF for restoration usage is FFS. One possible way is carrying this information with the 401 challenge message, or with the 200 OK message. Or after the registration procedure succeeds, the S-CSCF then sends a NOTIFY message to carry the second P-CSCF information.

NOTE 5: Sending a NOTIFY that does not have information about the SUBSCRIBE that originated that NOTIFY was considered as a potential door to a Denial of Service attack by SA3. This could be avoided if the operators could ensure their underlying access network is secure, e.g., by enabling the GPRS or UMTS encryption, or by activating IMS confidentiality protection.

6.4 Monitoring P-CSCF Health

6.4.1 Introduction

This solution is based in the assumption that it would be possible to notify the UE about a service interruption of the P-CSCF. When this notification is received, it should be possible for the UE to perform e.g. an initial registration and return to normal operation rapidly, without the need to wait for an originating SIP request from the UE to detect the problem, and making the chances of losing any terminating traffic very small. Two alternatives for the implementation of this notification are described in clauses 6.4.2 and 6.4.3.

6.4.2 Monitoring P-CSCF health from the UE

For unreliable transport (UDP, DCCP) STUN keep-alive can be used; this is possible when STUN server is in P-CSCF. This solution is considered in IETF Draft draft-ietf-sip-outbound ref [7]. The UE must be aware that STUN keep-alive is used for the restoration procedure; short keep-alive timer is to be used when the keep-alive mechanism is used for the restoration. In addition to that, for this solution to be effective, the STUN server needs to be within the P-CSCF. For reliable transport (TCP, SCTP) the TCP keep-alive and/or the double CRLF method could be used. In this last case, the solution will be only partially effective, since the TCP Layer maybe working even if the P-CSCF application using that protocol layer fails to respond.
This kind of solution has the problem of battery drainage of the UE for mobile access networks.

6.4.3 Monitoring P-CSCF health from the IP GW

It is possible for the IP-GW (IP-GW is the IP connection point in the IP-CAN, that would be GGSN in GPRS, PDN-GW in EPC, PDG for I-WLAN and potentially other network elements in other IP-CANs) to monitor the health of the P-CSCF (e.g. through the Gi Interface in the GPRS case, see 3GPP TS 29.061 [8]). This monitoring could also include some additional explicit signalling with an indication of a P-CSCF restart. On the event of P-CSCF unavailability, the IP-GW could remove the corresponding P-CSCF address from the list of P-CSCF addresses that are sent to the UE. On the event of a P-CSCF restart, the IP-GW could send an indication to the UEs with the list of P-CSCF addresses that are available to initiate a new registration. The following figure shows this flow:

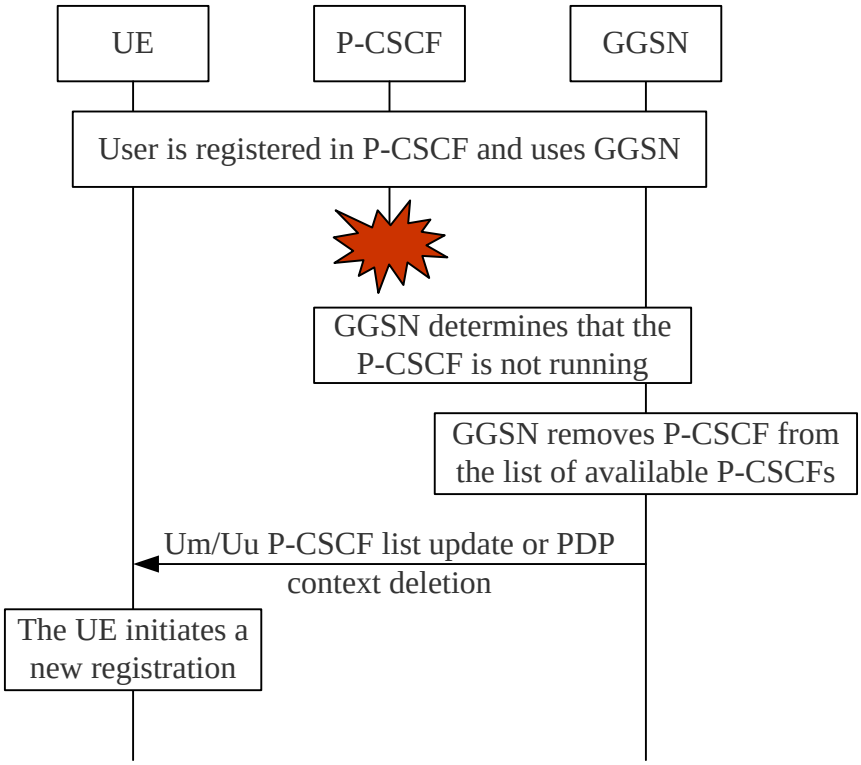

Figure 6.4.3.1. Notification of P-CSCF restart to UE from IP-GW (in this example the GGSN)

The updated list of P-CSCF addresses would be sent by GGSN/P-GW to all UEs that were initially provided with the unavailable/restarted P-CSCF in Protocol Configuration Options (PCO, see 3GPP TS 24.008, subclause 10.5.6.3) at CreatePDPContextResponse/Create Bearer Request (see 3GPP TS 29.060 and 3GPP TS 29.274), even if the list includes P-CSCF addresses that are currently being used (i.e. some UEs did not select the unavailable P-CSCF at initial registration). In addition, this solution would not support scenarios where P-CSCF discovery is performed by a DHCP server (see 3GPP 29.061, subclause 13a.2.1), in which the GGSN acts as a relay agent. In such a case, the P-CSCF addresses are not pre-configured in the APN, so the GGSN would not know which P-CSCF addresses are possibly being used by each user, neither the updated list of P-CSCF addresses.
P-GW provides support for P-CSCF discovery by requesting and provisioning of P-CSCF address (es) within the PCO IE. It also provides support acting as a DCHP client towards an external DHCP server in the initial access procedures. Hence, the solution above would be valid for EPS IP-CAN if GTP is used.
It could be possible for the IP-GW to be aware of the P-CSCF that each UE contacted at initial registration, independently of the P-CSCF discovery mechanism (i.e. provided by the IP-GW, via DHCP server, or pre-configured in UE). The P-CSCF could inform the IP-GW, upon receiving initial registration, about UE's IP-address/P-CSCF address pair, via Rx (using AAR command, see 3GPP TS 29.214, section 5.6.1). The PCRF could use a PUSH procedure to send this information via Gx to IP-GW (using RAR command, see 3GPP TS 29.212, section 5.6.4). The following figure shows this flow:

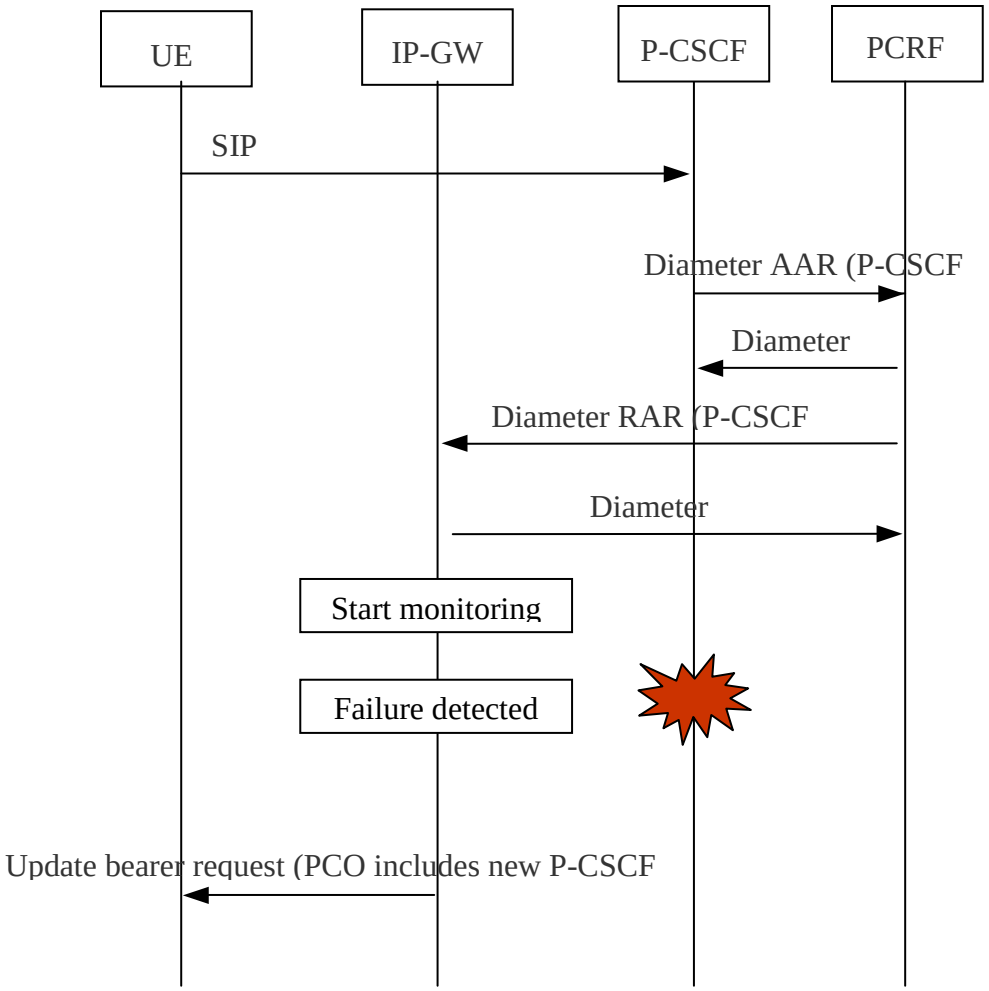

Figure 6.4.3.2. Notification of P-CSCF selected by UE via Gx/Rx

One alternative for the monitoring of health from the IP-GW to the P-CSCF, would be the implementation of an Echo Request similar to the one available in GTP between GGSN and SGSN (see 3GPP TS 29.060 [9], clause 7.2). This allows for the detection of unavailability (lack of response to an Echo Request) and a restart (an Echo Response with an incremented Restart Counter). An alternative method would be to monitor the incoming traffic from the P-CSCF and react when the P-CSCF doesn't send traffic in a specific period of time. This statistical monitoring would not be effective in low traffic scenarios.
There are also several alternatives for the notification that triggers the new registration by the UE. For the GPRS/EPS IP-CAN, the use of the PCO (see 3GPP TS 24.008, subclause 10.5.6.3) in Update PDP Context Request/Update bearer request (see 3GPP TS 29.060 and 3GPP TS 29.274) does not cause any side effects, but it needs to be supported also by the UE. It seems that the only way to trigger a new registration from the GGSN without adding new requirements to the UE, is to delete the PDP Context/bearer used for IMS SIP signalling. This would be the only alternative if a DHCP server is used for P-CSCF discovery. If a dedicated bearer is used for IMS, dropping the bearer could be the suggested proposal. Similarly in other IP-CANs, if no specific message is available to trigger a new IMS registration in the UE, the last resort can be to remove IP connectivity forcing a new IP connection and a new IMS registration. Note that in case of a bearer dedicated for multiple applications (in addition to IMS), this alternative would impact the rest of applications.

If there are existing SIP sessions established, the IP-GW could inform the PCRF about the loss of bearer to avoid relying on the UE sending a SIP BYE or initial SIP REGISTER. This would remove the sessions in PCRF, since they are not valid any longer.
.

6.5 Possible Solution for SIP-AS Service Restoration

One possible solution is to extend the initial Filter Criteria to contain more than one SIP-AS name or address with pre-configured priority order respectively. The S-CSCF is able to select one of the SIP Application Servers based on the matched Filter Criteria for the service request. When the selected high priority SIP-AS does not respond the service request or returns a failure response to the S-CSCF indicating no available resource, the S-CSCF may re-select a low priority SIP-AS, which could provide the same service to the user based on the matched Filter Criteria. The timer length and the re-selection times could be specified by the operator or the user. The restoration is shown as the following:

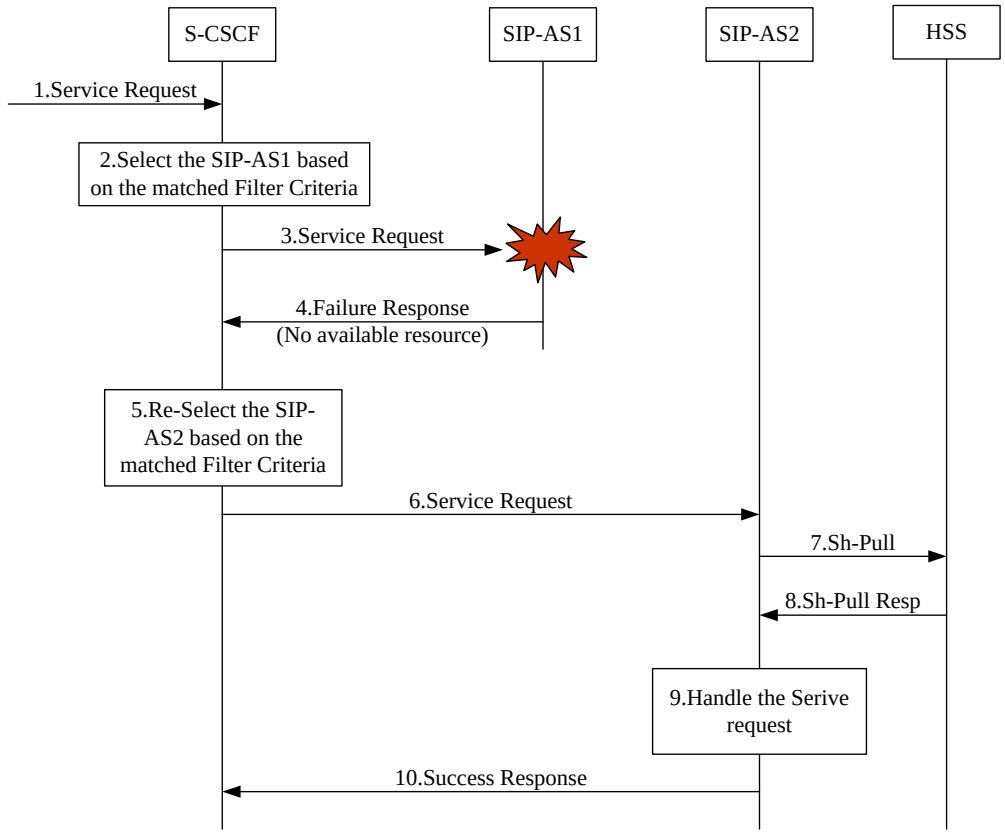

Figure 6.5.1 Re-select the SIP-AS for the User if the Pre-Selected SIP-AS Failed

1. The S-CSCF assigned to the user receives a service request (the third party registration request, or the originating/terminating service request).
2. The S-CSCF checks the initial Filter Criteria and selects the highest priority SIP-AS1 in the matched Filter Criteria to provide services to the user.
3. The S-CSCF sends the service request to the selected SIP-AS1.
4. If the SIP-AS1 is in failure, the SIP-AS1 may return a failure response indicating no available resource, or the SIP-AS1 does not respond the service request until the timer in the S-CSCF turns out.
5. The S-CSCF re-selects a new SIP-AS configured in the matched Filter Criteria to provide services to the user.
6. The S-CSCF sends the service request to the re-selected SIP-AS2.
7. The SIP-AS2 may send Sh-Pull to the HSS to query shared service data if the SIP-AS2.
8. The HSS send shared service data to the SIP-AS2 in the Sh-Pull Resp message.
9. The service request is handled by the SIP-AS2.
10. The SIP-AS2 may respond a success response to the S-CSCF to inform that the service request is successfully handled by the SIP-AS2.

Note 1: The principle supporting this alternative would require agreement in SA2.

6.6 Update of S-CSCF Name in the HSS after Loss of Data

6.6.1 Introduction

This chapter focuses in the recovery of temporary data, since the HSS is assumed to have mechanisms that ensure data persistency of permanent data. For implementations in which temporary data is not volatile, the problem scenario described in clause 5.7 does not take place, and this solution will not be needed. The temporary information that will pose the most difficulties for processing traffic once lost is the S-CSCF name. As indicated in clause 5.7.4, loss of the S-CSCF name could cause double S-CSCF assignment and processing of SIP requests as unregistered when the user is in registered state. There are two ways in which the S-CSCF name can be restored in the HSS:

- With a request from the HSS when it resumes operation. The HSS could send a Diameter request similar to the MAP RESET so that all S-CSCFs refresh their state for all users. This needs to be done in a gradual way so that it does not create a burst of traffic in the network.
- Adding some procedures to find out which is the S-CSCF serving the user when processing each SIP request for a user potentially affected by the loss of data in the HSS. This will also create some additional traffic, which might be larger in volume (it implies searching in all possible S-CSCFs), but it is distributed in time.

It should also be possible to combine both alternatives (the Diameter Restart indication could be sent so that the S-CSCF name is updated in the HSS by the S-CSCF as soon as possible, and still for those users whose S-CSCF name hasn't been updated the procedure to find out the S-CSCF is triggered). The following clauses provide a description of these two options. Clause 6.6.2 describes the use of the Diameter Restart indication, clause 6.6.3 describes the handling in the HSS of a new flag to indicate when to search for the S-CSCF serving the user and clause 6.6.4 describes the searching procedure.

6.6.2 Restart Indication

The principle is very similar to that of the MAP RESET indication. The HSS will send a message to all configured S-CSCFs, and these will mark all users from that HSS as "Not Confirmed in HSS". For all users marked as "Not Confirmed in the HSS", the S-CSCF will update the S-CSCF name in the HSS upon any activity for that user.
This solution alone does not cover the problem in clause 5.7.4, which was the most serious case. Since terminating SIP requests and REGISTER requests go through the I-CSCF before reaching the S-CSCF, they may still go to the wrong S-CSCF. A change in the procedure so that the S-CSCF does not wait for any activity for the users and refreshes the S-CSCF name immediately will create a burst of massive traffic in the network and still allow for a considerable window of time in which the problem could happen.

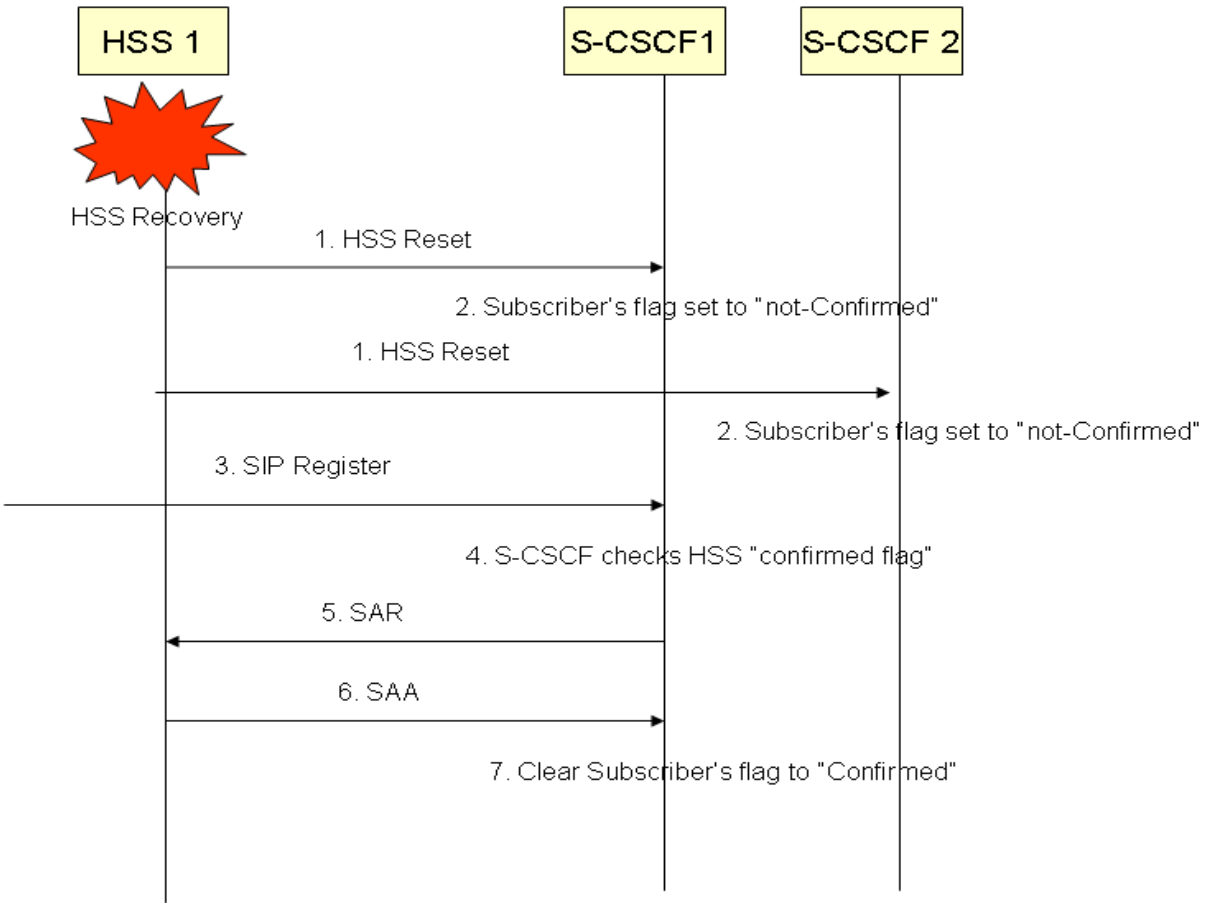

Figure 6.6.2.1. HSS-Reset in mobile originated session case

1. Upon completion of HSS recovery procedures, the HSS1 broadcasts the HSS-Reset message which is sent to S-CSCF1 and S-CSCF2.
2. S-CSCF1 and S-CSCF2 receive the reset message and set the “HSS Confirmed” flag against each subscriber to “not confirmed”.
3. User tries to register with IMS network.
4. Since this is an integrity protected register message, the S-CSCF checks the status of the “Confirmed flag”.
5. S-CSCF initiates a Server Assignment Request (SAR) towards the HSS1.
6. HSS1 sends a Server Assignment Answer(SAA).
7. Upon receiving a SAA with a valid user profile, the S-CSCF then clears the “HSS Confirmed” flag to “confirmed”.

6.6.3 S-CSCF Name Check Required Flag

The procedures that search for the correct S-CSCF, are based on the knowledge of when this search should be performed. The proposal is to store this in the form of a flag in the HSS. The HSS will set this flag for all users after the restart, and the flag will be cleared when the S-CSCF name is set for the user or after a period of time that should be a bit larger than the re-registration timer in the S-CSCF.

6.6.4 S-CSCF Name Check by the HSS

This means that for all users with the S-CSCF Name Check Required flag, the HSS will send a request to all S-CSCFs where the user could be located before answering a UAR request from the I-CSCF. If one of the S-CSCFs was serving that identity or any other identity of the same IMS subscription, the concerned S-CSCF will respond indicating this. The HSS will then send a response to the I-CSCF with the name of the S-CSCF. If none of the S-CSCFs was serving that identity or any other identity of the user or any user of the same IMS subscription, the HSS will also receive an indication and it will send the capabilities to the I-CSCF so that a new S-CSCF is selected.
In order to avoid double allocation of S-CSCF if the identity in the SIP request is Not Registered, but it belongs to an IMS Subscription that has other identities in Unregistered or Registered State, the S-CSCF must have knowledge of the relationship between all Private User Identities in the IMS Subscription (e.g. by mandatory use of Associated- Identities AVP). The S-CSCF will be required to check if any of those Private User Identities are being handled (and not just the

Private User Identities with an associated Public User Identity that is in Registered or Unregistered State), and respond with the corresponding indication to the HSS.
The following figure shows an example traffic flow with this proposal:

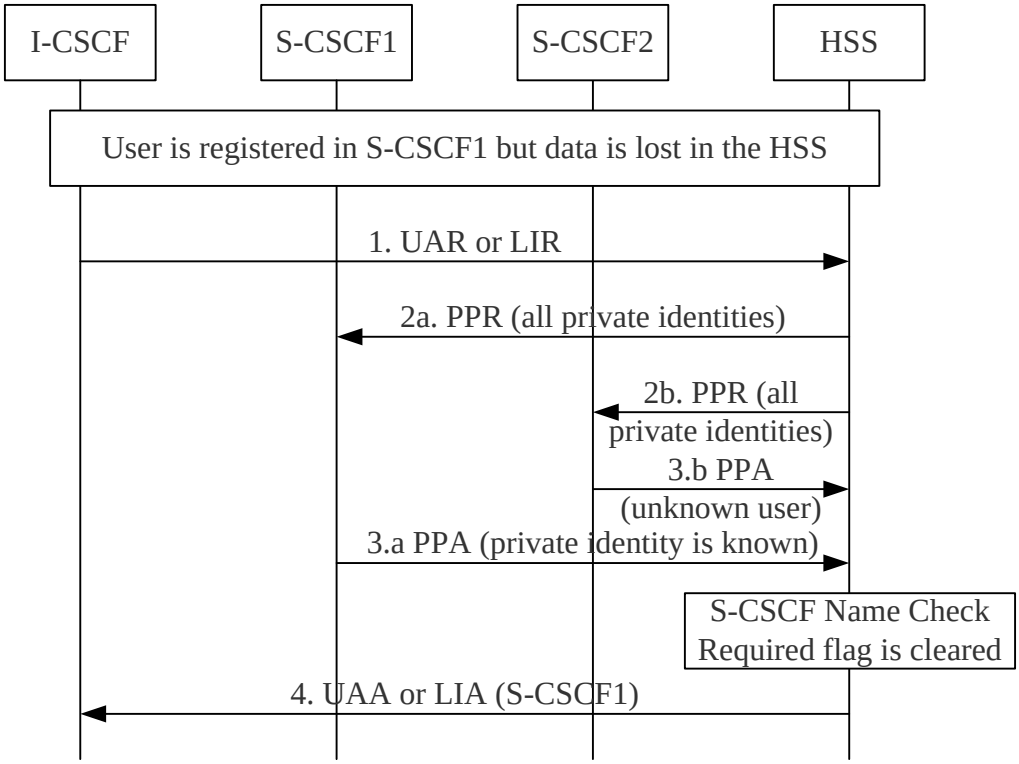

Figure 6.6.4.1. S-CSCF Name Check by the HSS

The figure proposes the use of PPR to locate the S-CSCF. This will also update the profile in the S-CSCF with whatever is stored in the HSS (which should be done before processing any requests). To allow this procedure it is required to have a list of all possible S-CSCF names in the HSS.

NOTE: This procedure requires enhancements to the S-CSCF (additional handling in the PPR) and the HSS (list of all possible S-CSCF names, procedure to send the PPR to those S-CSCFs and handling of S-CSCF name check required flag).

6.7 Forking Service Restoration

For forking service restoration, the S-CSCF information defined in the section 6.1.3 that the HSS sends to the S-CSCF could include the backup data associated with all the Private User Identities through which the restored Public User Identity has been registered. This could be done during the registration process with an additional information element in the SAR request, in addition to the basic set of information required to handle traffic.

During registration procedures, the HSS could send all the registered Private User Identities sharing the same Public User Identity which is being registered in the SAA, in addition to the basic user data to the S-CSCF. Then the S-CSCF compares the registered Private User Identities received from the HSS with the ones it stores. If there are any registered Private User Identities the S-CSCF does not have their registration data, the S-CSCF sends SAR with RESTORE indication to the HSS to retrieve the backup data for the registered Public User Identity, just as the S-CSCF does for terminating service restoration in the section 6.1.3. For this it is required to enhance the HSS and the S-CSCF. The changes to the protocol would be in the form of an additional information element in table 6.1.2.1 of 3GPP TS 29.228[10]:

Associated Registered Private Identities (See 7.X)	Associated-Registered-Identities	C	This AVP contains all Private Identities that are registered with the Public Identity received in the SAR command. The HSS shall send this information element if it implements the IMS Restoration Procedures and the value of Server-Assignment-Type in the request is REGISTRATION or RE_REGISTRATION and there are other Private Identities different from the Private Identity received in the SAR command being registered with the Public Identity received in the SAR command. If there are no other Private Identities different from the Private Identity received in the SAR command being registered with the Public Identity received in the SAR command, this AVP shall not be present.

6.8 Possible Solutions for SIP-AS Service Restoration

One possible solution is to delegate the routing of SIP requests post filter criterion matching to a representative AS (Rep-AS). The name of the representative AS is stored as part of the IFCs at the S-CSCF. On receiving a failure response from a selected SIP-AS or in the event of a SIP request timeout, the Rep-AS selects the next most suitable AS on behalf of a public Identity.

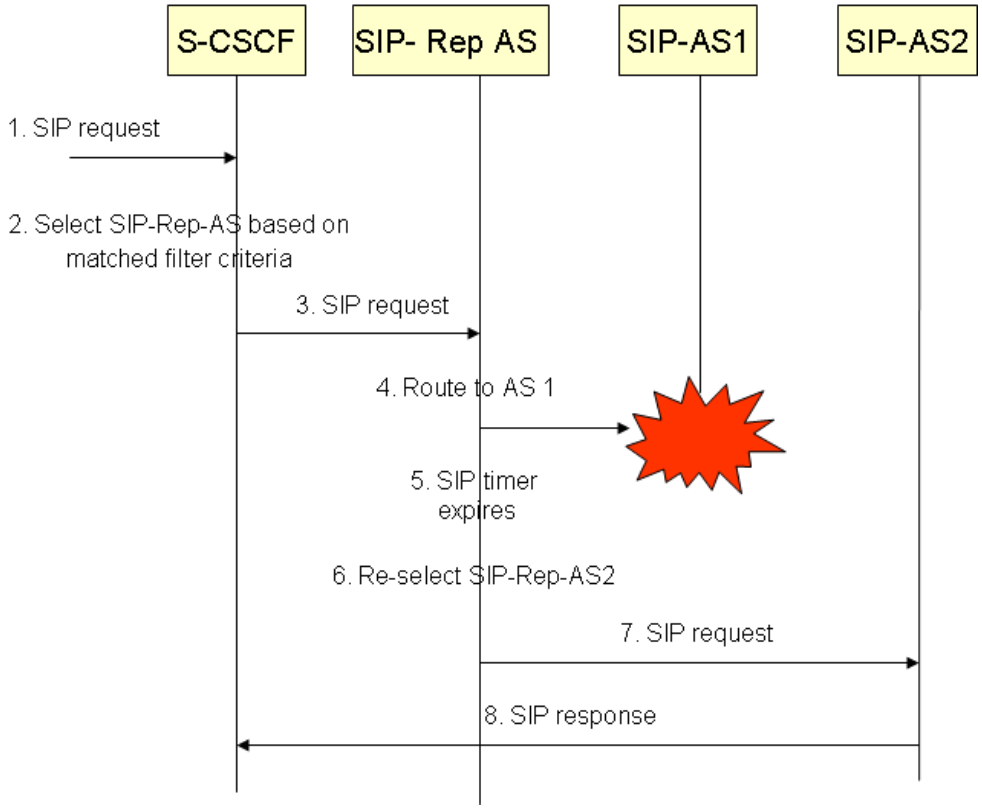

Figure 6.8.1. Re-selection of AS on behalf of a public identity by Rep-AS

1. The S-CSCF assigned to the user receives a service request (the third party registration request, or the originating/terminating service request).
2. The S-CSCF checks the initial Filter Criteria and selects the Rep-AS in the matched Filter Criteria to provide services to the public identity.
3. The S-CSCF sends the service request to the selected SIP Rep-AS.
4. The Rep-AS forwards the request to the SIP-AS1 based on received service request.
5. If the SIP-AS1 is in failure, the SIP-AS1 may return a failure response indicating no available resource, or the SIP-AS1 does not respond the service request until the timer in the S-CSCF/Rep-AS runs out.
6. The Rep-AS re-selects a new SIP-AS to provide services to the user.
7. The Rep-AS sends the service request to the re-selected SIP-AS2.
8. The SIP-AS2 may respond a success response to the S-CSCF (directly, via the Rep-AS) to inform that the service request is successfully handled by the SIP-AS2.

NOTE: For signalling efficiency subsequent messages may be routed directly from S-CSCF to SIP-AS2.

Another possible solution is that upon failure detection, the S-CSCF queries a specific server (DNS, SLF, etc.) to retrieve a list of Application Server that may handle the request. The query could be based on a specific info in the iFC (e.g. service name/AS name). The S-CSCF forwards the request to the AS, which then contacts the HSS to retrieve service info for the given IMPU over the Sh interface.

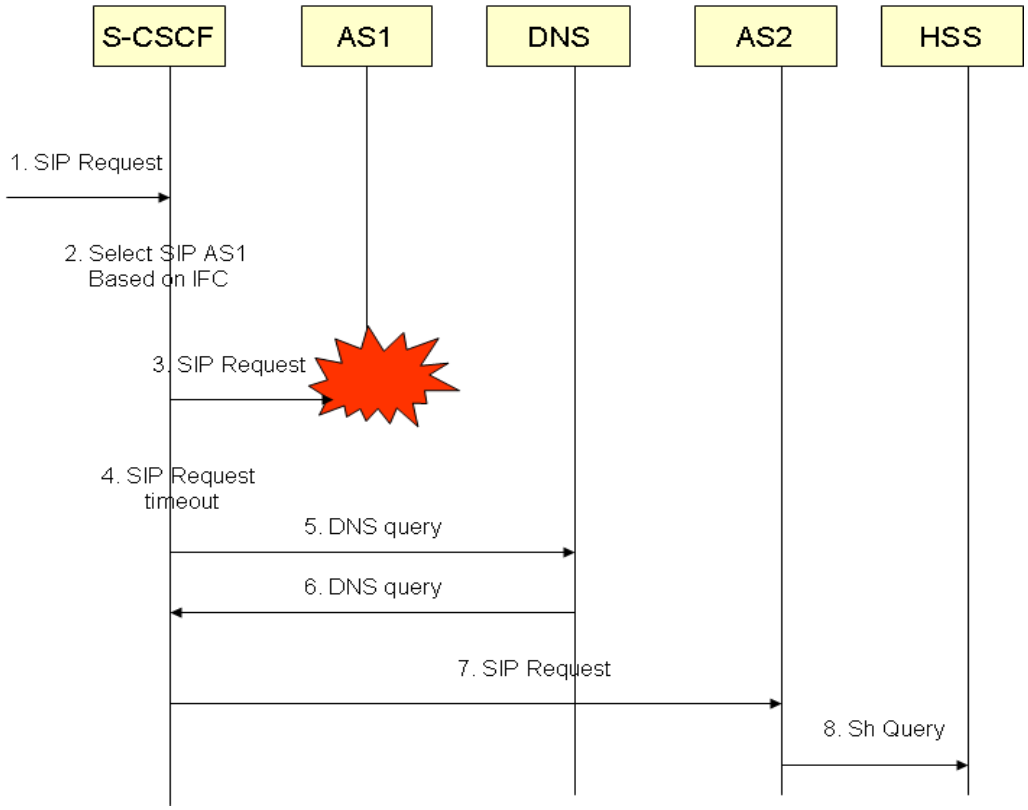

Figure 6.8.2. Re-selection of AS by DNS query

6.9 AS Behaviour After HSS Recovery

Application servers are able to store transparent data and non-transparent data at the HSS. Additionally the AS is allowed to subscribe to be notified of changes to the data using the Sh Interface. If the subscriptions to notifications are lost, the SIP-AS will not receive the notifications that it expects until it subscribes again. (see section 5.7.5).
Once the HSS recovers from a failure, it sends a Reset message to all the Application Servers for which it has stored either transparent or non-transparent data. Upon receiving such a Reset message the AS may re-subscribe to be notified of changes. Additionally, the AS may trigger a user data read message to download a fresh instance of user data after it has received such a HSS-Reset message. These re-subscription or user data read messages may happen at a later point of time.

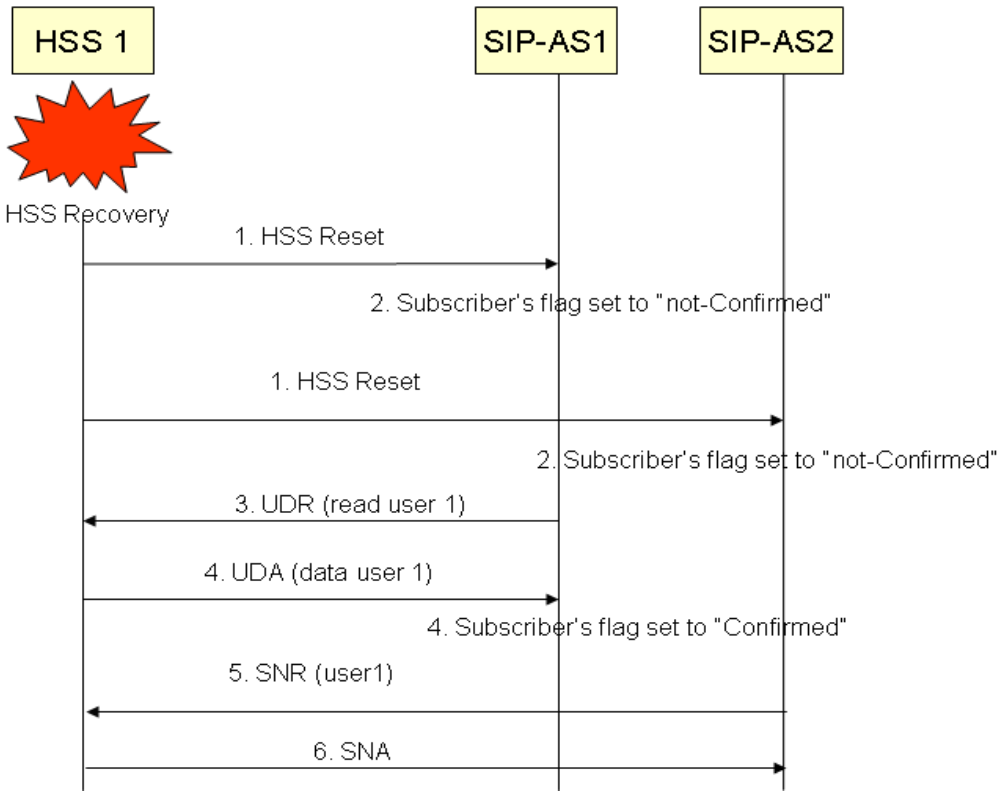

Figure 6.9.1. Application Server behaviour post HSS recovery

1. Upon completion of HSS recovery procedures, the HSS1 broadcasts the HSS-Reset message which is sent to AS1 and AS2.
2. AS1 and AS2 receive the reset message and set the “HSS Confirmed” flag against each subscriber to “not confirmed”.
3. AS1 initiates a user data read message to obtain the latest version of user data.
4. Upon successfully receiving the data , AS1 clears the “HSS confirmed” flag to “confirmed”.
5. AS2 initiates a subscription to notification on behalf of user 1.
6. HSS sends an acknowledgment confirming the re-subscriptions.

6.10 HSS Failover with no loss of service

6.10.1 Introduction

This chapter focuses on the procedures needed to guarantee service in networks where the HSS’s permanent and temporary data is replicated to a backup system and no loss of volatile data is experienced. In order for seamless service in networks deployed this way, the I-CSCF, S-CSCF and AS nodes need to support Diameter failover.

NOTE: The HSS data replication mechanism is outside the scope of this document.

6.10.2 Diameter FAILOVER

The peer connection principles and failover mechanisms for Diameter is described in IETF RFC 3588 [11], clauses 5.1 and 5.5.4 respectively and applies to the Cx, Dx, Sh and Dh interfaces. More specific procedures related to IMS elements are found below.

6.10.2.1 SLF

The SLF is queried over the Dx interface by the S-CSCF and the I-CSCF and over the Dh interface by the AS. In networks where the HSS is deployed with one or more backup systems, the SLF will return all HSS identities associated with the user in the query response. The order of the HSS identities in the query response is significant; i.e. the first identity in the list is the primary HSS, the second identity is the secondary HSS and so forth.

6.10.2.1 I-CSCF, S-CSCF and AS

If a network is using the SLF function, the I-CSCF, S-CSCF and AS will at a minimum be configured to address a primary SLF and a backup SLF. How the I/S-CSCF/AS may load balance across the SLFs is implementation specific. The I/S-CSCF/AS will use the primary HSS address returned by the SLF to communicate with the HSS for the duration of the session. Only if the I/S-CSCF/AS detects a failure in the communication with the HSS, will the I/S-CSCF/AS attempt to communicate with the secondary HSS.
If no SLF is deployed in the network, the I/S-CSCF/AS will have to be configured to allow for primary/backup HSS's for each user handled by the I/S-CSCF/AS.

7 Conclusions and recommendations

7.1 S-CSCF Service Interruption

This Technical Report recommends implementing a solution with the following principles:

- A specific error is returned to the UE in the event of lack of response from the corresponding S-CSCF or the S-CSCF does not have the user data (see clause 6.2.2.1 and 6.2.2.2). For the second case, the specific error could be sent to the UE immediately when the S-CSCF receives the originating request as indicated in clause 6.2.2.2 or after the S-CSCF fails to retrieve user data from the HSS by the procedure indicated in the clause 6.1.5.
- Modification of the SAR-SAA so that:
  - It is possible for the S-CSCF to store S-CSCF specific information (such as that described in clause 6.1.2) in the HSS.
  - Modification of the UAR-UAA and LIR-LIA so that the I-CSCF is able to assign a new S-CSCF upon S-CSCF failure detected during registrations (see clause 6.1.6), processing of terminating SIP request and processing of originating SIP requests from a SIP-AS. S-CSCF failure may be detected in I-CSCF, SIP-AS or P-CSCF. S-CSCF reassignment is always performed by the I-CSCF.
  - A registration will be triggered from the P-CSCF if S-CSCF failure is detected when processing an originating SIP request from the UE.
  - The HSS will allow S-CSCF name overwriting after an explicit request for capabilities from the I-CSCF.
  - If the S-CSCF receives a SIP request for a user that it does not recognize, it will send an SAR to the HSS to check whether there is anything stored for that user there before sending a reply (see clause 6.1.5).
  - The HSS will send all the registered Private User Identities sharing the same Public User Identity which is being registered in the SAA during registration procedure (see clause 6.7).

7.2 S-CSCF Re-Selection during Re-Registration

The procedures described in clause 6.1.6 has been selected as the preferred way to handle the scenario for S-CSCF service interruption detected by the I-CSCF during processing of re-registration request. It is recommended to proceed with the specification of a procedure for the I-CSCF in which a new S-CSCF is re-selected in the event of lack of response from the corresponding S-CSCF. This will require changes to 3GPP TS 29.228 and TS 24.229, so contributions need to be submitted to CT4 and CT1 in order to complete this task.

7.3 SIP-AS Service Interruption

This TR recommends that an architecture similar to the one described in clause 6.8 when required in order to solve the problems described in clause 5.5. No further specification work seems to be required.

7.4 HSS Service Interruption

This Technical Report recommends that in order to avoid the problem scenarios in clause 5.7 the HSS should be implemented in a way that critical information can be assumed to be always available. Permanent data and also some temporary data such as the Sh transparent data, S-CSCF name and IMS registration status of a user shall be considered critical information. Sh subscriptions to notifications may also be included in this list of critical information. This list of critical information may need to be described in normative text.

7.5 P-CSCF Service Interruption

This version of the Technical Report concludes that it is feasible to finalize a full normative solution for the current release. It is recommended that if the UE needs to become aware of the P-CSCF service interruption, the use of a

procedure that informs the UE about the health of the P-CSCF could be used such as those listed in clause 6.4.2. For an IP-CAN such as GPRS, where due to battery consumption constrains the UE is not expected to support this kind of procedures, procedures such as monitoring of the P-CSCF from the IP-CAN as described in clause 6.4.3 could be used. This technical report recommends implementing a solution with the following principles:

- P-CSCF needs to inform the IP-GW via Rx/Gx about the UE's P-CSCF selection
- IP-GW monitors the health of the P-CSCF (via echo request/echo response, statistical monitoring is optional)
- IP-GW informs the UE (either with PCO or simply removing the IP-CAN)

This will require changes to 3GPP TS 23.380, TS 29.212, TS 29.213, TS 29.214, TS 24.008 and TS 24.229, so contributions need to be submitted to CT4, CT1 and CT3 in order to complete this task.

Annex A:
Change history

Change history							
Date	TSG #	TSG Doc.	CR	Rev	Subject/Comment	Old	New
2008-06	CT#40	CP-080244			Approved in CT#40	2.0.0	8.0.0
2008-12	CT#42				Copyright Notification updated	8.0.0	8.0.1
2009-09	CT#45	CP-090560	0001		Study for P-CSCF restoration procedures	8.0.1	9.0.0