NICC ND 1522 V2.1.1 (2020-09)

NICC Document

Report into implementation of Secure Telephone Identity Revisited (STIR) in the UK

NICC Standards Limited

c/o TWP ACCOUNTING LLP,
The Old Rectory,
Church Street,
Weybridge,
Surrey KT13 8DE

Tel.: +44(0) 20 7036 3636

Registered in England and Wales under number 6613589

Liability

Whilst every care has been taken in the preparation and publication of this document, neither NICC, nor any working group, committee, member, director, officer, agent, consultant or adviser of or to, or any person acting on behalf of NICC, nor any member of any such working group or committee, nor the companies, entities or organisations they represent, nor any other person contributing to the contents of this document (together the “Generators”) accepts liability for any loss or damage whatsoever which may arise from the use of or reliance on the information contained in this document or from any errors or omissions, typographical or otherwise in the contents.

Nothing in this document constitutes advice. Nor does the transmission, downloading or sending of this document create any contractual relationship. In particular no licence is granted under any intellectual property right (including trade and service mark rights) save for the above licence to download copy, store and distribute this document and to produce derived copyright works.

The liability and responsibility for implementations based on this document rests with the implementer, and not with any of the Generators. If you implement any of the contents of this document, you agree to indemnify and hold harmless each Generator in any jurisdiction against any claims and legal proceedings alleging that the use of the contents by you or on your behalf infringes any legal or other right of any of the Generators or any third party.

None of the Generators accepts any liability whatsoever for any direct, indirect or consequential loss or damage arising in any way from any use of or reliance on the contents of this document for any purpose.

The NICC Standards Web site contains the definitive information on the IPR Policy and Anti-trust Compliance Policy

Contents

Intellectual Property Rights....4

Foreword....4

Introduction....4

1 Scope....5

2 References....5

3 Definitions and abbreviations....5
3.1 Definitions....5
3.2 Abbreviations....7

4 Background to CLI generation....8

5 STIR Concept....10

6 Overview of STIR....13
6.1 STIR/SHAKEN Architecture....13
6.2 Typical Call Flow....15
6.3 Terminating Network actions....17

7 Distribution of credentials....19

8 International calls....21
8.1 Inbound international calls....21
8.2 Outbound international calls....22

9 Benefits and gap analysis of STIR....23

10 Implications for Network Protocols....28
10.1 UK SIP support for STIR PASSporT....28
10.2 Compact vs full form....28
10.3 PASSporT URI....28
10.4 Other considerations....29

11 Implementing the STIR Functions....29
11.1 Communication Provider Functions....29
11.2 Central Functions – Certificate Authority....30

12. Alternatives to STIR....30

13 Conclusions and Next Steps....31

Annex A (informative): Implementation Phases....33

Annex B: (Informative) Changes to call flow to support final stage implementation....37

Annex C (informative): Approaches considered for credential distribution....40

Annex D (informative): Bibliography....48

History....49

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to NICC. Pursuant to the NICC IPR Policy, no investigation, including IPR searches, has been carried out by NICC. No guarantee can be given as to the existence of other IPRs which are, or may be, or may become, essential to the present document.

Foreword

This NICC Document (ND) has been produced by NICC N-CLI TG; it builds upon a previous version, in light of subsequent experience of implementation in other countries.

Introduction

UK consumers receive large volumes of unsolicited and unwelcome marketing calls. Although some calls display a valid, reliable and authentic Calling Line Identity (CLI), in a significant proportion of cases the CLI is a spoof one which does not belong to the caller, and is included for display purposes solely to give the illusion of a legitimate call.

When CLI facilities were introduced, the population of the CLI information was carried out by the originating network, and with a very limited number of originating networks, the system was largely secure. However, over time the facility for callers to populate the Presentation Number CLI has been introduced, and the number of originating networks has dramatically increased. Both of these developments are to be welcomed – caller population of CLI allows a more meaningful number to be displayed, and more originating networks foster greater competition – but a side effect of this has been the loss of security, hence trust, of CLIs.

There are some measures that networks can take without needing modification to signalling systems; for example international calls which are fraudulently passed off as having been originated from a UK connection can be blocked at international gateways. However, there remain nuisance calls that are originated by bad actors within the UK, using CLIs without authorisation.

The Internet Engineering Task Force (IETF) has standardised a mechanism to digitally sign identities (such as CLIs) in order that terminating networks or endpoints can validate who populated the information, called Secure Telephony Identity Revisited (STIR). This document provides insight into the implementation issues should it be decided to adopt STIR technology in UK networks.

1 Scope

This document describes the benefits and implementation issues of adopting STIR technology to digitally sign CLIs in UK networks. The document provides a background to how UK CLIs are populated, and describes how STIR would interact with this. It describes how the STIR functions could be implemented in the UK, sets out the benefits of doing so, and identifies the remaining limitations of this mechanism in eliminating nuisance calls.

2 References

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For non-specific references, the latest edition of the referenced document (including any amendments) applies.

[1] ITU-T Recommendation E.164 (11/2010) “The international public telecommunication numbering plan”

[2] Communications Act 2003; http://www.legislation.gov.uk/ukpga/2003/21/contents

[3] ND1016: Requirements on Communications Providers in relation to Customer Line Identification display services and other related services

[4] ND1035: SIP Network to Network Interface Signalling

[5] RFC8224 (February 2018) “Authenticated Identity Management in the Session Initiation Protocol (SIP)”

[6] ATIS-1000074 (5th January 2017): "Joint ATIS/SIP Forum Standard – Signature-based Handling of Asserted information using toKENs (SHAKEN)"

[7] IETF RFC8225: “PASSportT: Personal Assertion Token”

[8] ND1034:UK SIPconnect Endorsement

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply:

Authentication Service: A STIR function described in Section 6.1 of this document.

Border Gateway: The node providing a Network-Network Interface to other service provider networks.

Calling Line Identity: A telephone number representing the calling party. The CLI may be a Network Number or a Presentation Number.

Certificate Authority: A STIR function described in Section 7 of this document.

Certificate Store: A STIR function described in Section 6.1 of this document.

Final Stage: As set out in Section 5 of this document, the implementation stage of UK STIR which would deliver Full Attestation.

Final Stage Variant: One of the options for the Final Stage as set out in Section 4 of this document.

Full Attestation: As set out in Section 5 of this document, the status where the entity signing the CLI can unequivocally assert that the originator of the call has rights to use both the Network Number and Presentation Number CLIs.

Gateway Attestation: As set out in Section 5 of this document, the attestation status used where a gateway operator, for example the operator of an international inbound gateway, has signed the CLI.

Interim Stage: An implementation of UK STIR whereby the Network Number would be signed, potentially on a Partial Attestation basis, as set out in Section 5 of this document, and the availability of a numbering database would allow terminating networks to assess whether the Network Number was valid for the originating network.

Key Store: A STIR function described in Section 6.1 of this document.

Network Number: The digits that comprise a unique E.164 [1] number that unambiguously identifies the point of ingress of the call to a Public Electronic Communications Network.

Network Termination Point: The physical point at which a Subscriber is provided with access to a Public Electronic Communications Network and which may consist of one or more lines.

Originating Call Server: A generic term to represent the element carrying out call session control functions for origination to the Public Electronic Communications Network. The Call Server functionality could be discrete or combined with other functions, for example border gateway functionality within a session border controller.

Partial Attestation: As set out in Section 5 of this document, the status where the entity signing the Network Number CLI can assert that it hosts the customer using a Presentation Number CLI, but cannot assert that the customer necessarily has the rights to use that Presentation Number CLI.

Presentation Number: A number nominated or provided by a subscriber to be used for display purposes and can be used to make a return or subsequent call.

Public Electronic Communications Network: Public network as defined in the Communications Act 2003 [2].

Start-up Stage: An initial implementation of UK STIR whereby the Network Number would be signed, potentially on a Partial Attestation basis, as set out in Section 5 of this document.

SIP Terminal: The terminal making or receiving a phone call, as described in Section 6.1. Note that for the purposes of this document the term is used generically, and encompasses, for example, functionality within an analogue telephony adaptor.

Terminating Call Server: A generic term to represent the element carrying out call session control functions for termination from the Public Electronic Communications Network. The Call Server functionality could be discrete or combined with other functions, for example border gateway functionality within a session border controller.

Treatment Policy Server: A STIR function described in Section 6.1 of this document.

Verification Service: A STIR function described in Section 6.1 of this document.

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

C7	Common Channel signalling system number 7 (used in legacy telephone networks)
CLI	Calling Line Identity
CP	Communications Provider
CSCF	Call Session Control Function
CVT	Call Validation Treatment
HTTPS	HyperText Transfer Protocol Secure
IBCF	Interconnection Border Control Function
IETF	Internet Engineering Task Force
IP	Internet Protocol
MoU	Memorandum of Understanding
NN	Network Number
P-A-ID	P-Asserted- Identity
PASSporT	Personal ASSertion Token
PBX	Private Branch eXchange
PN	Presentation Number
SHAKEN	Signature-based HAndling of asserted information toKENs
SIP	Session Initiation Protocol
SIP UA	SIP User Agent
SKS	Secure Key Service
STI – AS	Secure Telephone Identity Authentication Service
STI – VS	Secure Telephone Identity Verification Service
STIR	Secure Telephone Identification Revisited
URI	Uniform Resource Identifier

4 Background to CLI generation

As described in ND1016 [3], in the UK, two CLIs are conveyed on calls, namely the Network Number (NN) and Presentation Number (PN). The Network Number unambiguously identifies the point of entry into the public telephone network, whereas the Presentation Number is used for display purposes. According to UK SIP standard ND1035 [4] these are carried in the P-Asserted Identity (P-A-ID) and From header fields respectively; NICC Standards' research indicates this approach is adopted widely internationally, but not universally.

Note that in legacy C7 TDM signalling systems the Presentation Number is optional, but given STIR demands end-end IP, by the time of implementation of STIR there will always be two CLIs, albeit these could be the same number.

There are a large number of scenarios of how CLIs are generated in the context of SIP. Table 4.1.a below sets out the cases which cover the majority of call volumes.

Table 4.1.a: CLI scenarios in UK

Scenario	Network Number	Presentation Number
1. Residential line	Is populated by the originating network, representing the Network Termination Point from where the call originated	Theoretically there is no Presentation Number. However, as SIP mandates that the From header field is populated, and this is the field which contains the Presentation Number, the originating network populates both fields with the Network Number (i.e. the Presentation Number is a copy of the Network Number)
2. Business line where the customer has not requested a Presentation Number		
3. Business line with a static Presentation Number – this is known as a "Type 1" CLI	Is populated by the originating network, representing the Network Termination Point from where the call originated	Is populated by the originating network according to the instructions of the customer (and is the same for all calls)
4. Business line where the originating network has verified the number received from the customer – this is known as a "Type 2" CLI	Is populated by the originating network according to a static number, modified by CLI digits received from the originating customer on a per call basis (these having been verified as belonging to the customer); this uniquely represents the Network Termination Point from where the call originated	Is populated by the originating network according to a static number, modified by CLI digits received from the originating customer on a per call basis (these having been verified as belonging to the customer); this uniquely represents the Network Termination Point from where the call originated
5. Business line being entry into the public network from an enterprise network having multiple sites and/or extensions – this is known as a "Type 3" CLI	Is populated by the originating network, representing the Network Termination Point from where the call originated	Is populated by the originating network according to CLI information provided by the enterprise customer. No validation is carried out on a call-by-call basis, however the originating network must enter into an agreement with the customer that they will send only CLIs that they are authorised to use

Scenario	Network Number	Presentation Number
6. Business line being entry into the public network from an enterprise network that allows its users to dial into its network then make a breakout call – this is a form of “Type 4” CLI	Is populated by the originating network, representing the Network Termination Point from where the call was passed back into the public network	Is populated by the originating network according to CLI information provided by the enterprise customer which in turn should have copied across the Presentation Number details from the inbound leg. No validation is carried out on a call-by-call basis, however the originating network must enter into an agreement with the customer that they will send only CLIs that are received on the inbound leg.
7. Business line being entry into the public network from a call-centre that wishes to use a different CLI according to the customer campaign that’s being supported – this is known as a “Type 5” CLI	Is populated by the originating network, representing the Network Termination Point from where the call originated	Is populated by the originating network according to CLI information provided by the enterprise customer. No validation is carried out on a call-by-call basis, however the originating network must enter into an agreement with the customer that they will send only CLIs that they are authorised to use.
8. Calls being diverted by the public network	Network Number is passed unaltered as received on the inbound leg	Presentation Number is passed unaltered as received on the inbound leg
9. Calls being diverted by customer PBXs where it is wished that the original caller’s number be displayed – this is a form of “Type 4” CLI	Is populated by the originating network, representing the Network Termination Point from where the call was passed back into the public network	Is populated by the originating network according to CLI information provided by the enterprise customer which in turn should have copied across the Presentation Number details from the inbound leg. No validation is carried out on a call-by-call basis, however the originating network must enter into an agreement with the customer that they will send only CLIs that are received on the inbound leg.
10. UK mobile calling from their home network (i.e. not roaming)	Is populated by the originating network, representing the account associated with the mobile subscription	Theoretically there is no Presentation Number. However, as SIP mandates that the From header field is populated, and this is the field which contains the Presentation Number, the originating network populates both fields with the Network Number (i.e. the Presentation Number is a copy of the Network Number). Note: enterprise mobile customers may use Presentation Numbers, in which case the Type 1 Presentation Number row above should be referred to.
11. UK mobile roaming overseas (with home network routeing enabled)		
12. UK mobile roaming overseas (with direct routeing enabled)		
13. Foreign mobile roaming in UK	Is populated by the originating network, representing the account associated with the mobile subscription provided by the home network (i.e. will be from the country code of the mobile in question)	
14. Overseas call-centre wishing to display UK number: connected to UK network (e.g. long lined)	Is populated by the originating UK network, representing the Network Termination Point from where the call originated	Is populated by the originating network according to CLI information provided by the enterprise customer. No validation is carried out on a call-by-call basis, however the originating network must enter into an agreement with the customer that will send only CLIs that they are authorised to use

Scenario	Network Number	Presentation Number
15. Overseas call-centre hosted on overseas network wishing to display UK number: connected to overseas network	Is populated by the originating network, representing the Network Termination Point from where the call originated, i.e. should contain the country code of the host country	Is populated by the originating network according to CLI information provided by the enterprise customer. We can have no knowledge of what due diligence the originating network carries out

5 STIR Concept

The intent of STIR, as set out in RFC8224 [5], is that the originators of calls will digitally sign that they are authorised to use a given identity in order that at call termination this signature can be checked to validate the authenticity of the identity. This document examines the application of STIR to UK CLIs.

A long term solution could be that originating customers are able to sign their own CLIs, and the checking of signatures could be done by terminating customers themselves. However, the logistics of getting such a model running would be very complex, and a more practicable interim approach is that originating network operators carry out the signing process, and terminating operators carry out the checking process, providing some form of indication to terminating customers of the validity of the CLI. This does not preclude originating customers eventually signing their own CLIs – the phases subsequently presented in this report incorporate this – but it means that during the interim phase the number of entities involved in signing/checking CLIs is of the order of hundreds, rather than millions.

As it is the Presentation Number CLI which is displayed to end-users, at first sight it makes sense that this is the information that would be digitally signed via STIR. However:

1. Whereas each Network Number is used to identify only one location, the same Presentation Number can be used on multiple ingress points into the public network, potentially across multiple networks. If the originating network operator were to sign the Presentation Number, this would imply that the same number would be signed by multiple entities, i.e. multiple entities would need credentials to do so. This would complicate the verification of whether an originating network had the rights to sign the number (for example it might be necessary to have a database containing all the originating networks that a caller could use). In contrast, the Network Numbers follow a simpler assignment path from Ofcom directly to network operators, with the only complications being sub-allocations and portability (both of which are constrained within the network operator community).

2. There is a pioneer implementation of STIR in the USA, known as SHAKEN [6]. In contrast to the UK, the American approach to CLIs is to use the content of the P-A-ID header field for display purposes, and hence in the SHAKEN architecture it is the P-A-ID header field which is signed (it should be noted that this resulted in the issues identified in (1) with respect to being able to link a given number to an authorised originating network). Since SHAKEN signs the P-A-ID header field, following this approach in the UK is likely to mean easier adoption with equipment vendors, rather than taking a novel approach in a start-up phase.

For these reasons, it is recommended that STIR implementation in the UK initially be based upon signing of the Network Number by the originating network operator. This will confirm who accepted the call into the public network, and therefore who should be approached if there is any question-mark regarding the Presentation Number (or, to put it another way, signing and hence validating the Network Number provides a degree of confidence about the reliability of the Presentation Number). The remainder of this document terms this approach as the "Interim Stage".

The PASSport parameter in the SIP Identity header used in STIR/Shaken provides an attestation claim that relates to Originating Identity in the PASSport parameter. In the UK, the attestation claim within STIR will be used to allow the originating network operator to indicate whether they also populated the Presentation Number, potentially based upon information provided by the customer and authenticated by the originating network ("full attestation"), or alternatively that they have passed on unchanged a Presentation Number provided by the customer ("partial attestation"). Additionally, STIR allows operators of gateway facilities - such as international inbound gateways and interworking gateways from TDM networks - to indicate that the call has ingressed from a network not within the scope of the UK STIR implementation. In this case the Network Number would be signed with "gateway attestation". Table 5.a summarises the application of the attestation statuses. These attestation statuses would allow terminating networks to form a view as to the reliability of the CLI information, as set out in Section 6.3.

Table 5.a: Attestation Status

Status	Usage
Full	The CLI to be used for display purposes has been generated by the originating network itself, or if received from a customer, the originating network has checked that it is reliable and authentic for usage on the call
Partial	The originating network has received the CLI to be used for display purposes from a customer with which it has a Type 3/4/5 agreement
Gateway	A gateway (for example international) has received the call and signed the CLI contents to indicate that it admitted it to the UK public network

The proposed Interim Stage does not provide absolute authentication of all Presentation Numbers used for display, but it may be that this limited level of signing is sufficient to re-establish public confidence in CLI. If it is not, three options, termed Final Stage Variants for the remainder of this document, are foreseen for a long term solution. These Variants are not necessarily mutually exclusive – it is envisaged that the choice could be made according to which is most appropriate for the individual originating customer use case:

I. The originating network operator would verify the received Presentation Number against a whitelist of acceptable numbers for that customer, and only then be allowed to sign the Network Number (which would then be on a full attestation basis), or

II. The originating customer would sign the Presentation Number, and this signature would then be verified by the originating network in order to determine whether to sign the Network Number (i.e. no signing of the Network Number unless there had been a valid signed Presentation Number), or

III. The originating customer would sign the Presentation Number and this would be passed through to be verified at the terminating end.

At this Final Stage, it would be acceptable for a terminating network to reject calls with Network Number signed only on a partial attestation basis, unless they were accompanied by a signed Presentation Number.

In v1.1.1 of this report, it was recommended that the UK move directly to a STIR implementation that made use of a common database of numbers in order to confirm that the entity signing the CLI has the rights to use that number. However, in light of implementation in the USA (where, at the time of writing, the usage of such a database isn't mandated), this recommendation has been reviewed. Omitting a database check reduces the value of STIR in that such a solution confirms which network originated the call, but not whether they had the rights to do so with the CLI concerned. Conversely, it removes a significant implementation barrier, hence would ease start-up. For this reason, NICC now recommends that a Start-up Stage be incorporated, during which there is no database hence no check of the originating network's right to use the CLI.

This means that the recommended implementation of STIR in the UK is:

1. Start-up Stage: the originating network signs the Network Number, but with no database to confirm rights to use a number, terminating networks must take a view on the reliability of the CLI according to their experience of the originating network.
2. Interim Stage: the Network Number is signed by the originating network and can be used to give confidence in the Presentation Number. The rights of the originating network to sign the specific Network Number can be verified by reference to a numbering database.
3. Final Stage (if necessary): according to the Variant adopted, either originating network must check the Presentation Number with respect to a whitelist before signing the Network Number, or the Presentation Number itself is signed.

Within the proposed Start-up, Interim and Final stages, it would be necessary to have sub-phases in which STIR implementation was kick-started. For example, there is little merit in terminating networks seeking to verify signed CLIs prior to originating networks carrying out that signing function. Annex A describes how these sub-phases could work.

6 Overview of STIR

6.1 STIR/SHAKEN Architecture

Figure 6.1.a provides an overview of the STIR/SHAKEN architecture as applied to UK networks.

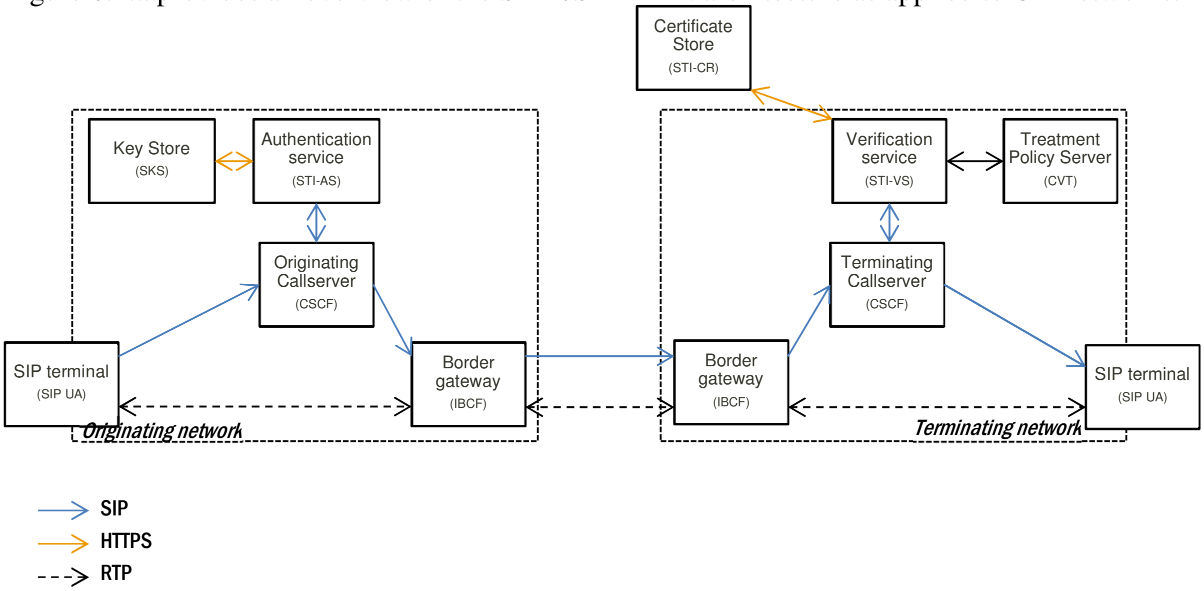

Figure 6.1.a: STIR/SHAKEN architecture

The functions in the architecture are as follows:

SIP Terminal (SHAKEN terminology – SIP User Agent, SIP UA). This is the terminal authenticated by the service provider network. This terminal could take a variety of forms, for example be a standalone SIP phone, be software on a computer, or be terminal adaptor functionality within a home router (presenting an analogue interface to the end user). Depending upon the nature of the SIP Terminal, it could be considered to be within the Originating network trust domain (e.g. if it was a terminal adaptor solely under the control of the Originating network), or be outside the Originating network (e.g. if it was a standalone terminal sourced from a third party). When the terminal is under direct management control of the telephone service provider, the service provider network can fully attest the CLI in originating SIP INVITE requests initiated by the terminal, otherwise it can do it only on a partial attestation basis unless it has some mechanism to verify the CLI received from the caller.

Originating Call Server (SHAKEN terminology – Call Session Control Function, CSCF). The Call Server is the SIP registrar and routeing function. It queries the Authentication Service with the CLI to be used on the call, in order that it can be signed. (In the UK interim stage application of STIR, the CLI to be signed is the Network Number, with the trust level for the Presentation Number being derived by implication from this).

Authentication Service (SHAKEN terminology – Secure Telephone Identity Authentication Service, STI-AS). The Authentication Service is a SIP application service that provides the function of authentication service described in RFC8224. It validates the CLI, queries the Key Store for the private key for that number, and digitally signs the P-A-ID header field.

Key Store (SHAKEN terminology – Secure Key Store, SKS). The Key Store is a highly secure element that contains the private keys accessed by the Authentication Service. Section 7 considers how this function would be populated.

Border Gateway (SHAKEN terminology – Interconnection Border Control Function, IBCF). The Border Gateway represents the Network-Network Interface (NNI) between service provider networks.

Terminating Call Server (SHAKEN terminology – Call Session Control Function, CSCF). The call server is the SIP registrar and routeing function. It queries the Verification Service with the signed CLI to determine call treatment.

Verification Service (SHAKEN terminology – Secure Telephone Identity Verification Service, STI-VS). The SIP application server that performs the function of the verification service defined in RFC8224. It has an interface to the Certificate Store that is referenced in the SIP Identity header field to retrieve the provider public key certificate.

Treatment Policy Server (SHAKEN terminology – Call Validation Treatment, CVT). The function that once the signature is positively or negatively verified, determines call treatment. For example it could on a per-network or per-subscription basis, accept or reject the call, or supply information that could be passed to the SIP terminal on the reliability and level of attestation of the CLI, that could be used to cause a specific display or ring tone.

Certificate Store (SHAKEN terminology – Secure Telephone Identity Certificate Repository). This represents the publicly accessible store for public key certificates. Section 7 considers where this function would practicably reside.

6.2 Typical Call Flow

Figure 6.2.a provides a typical call flow for the UK Start-up and Interim Stages (as described in Section 5), where CLIs are authenticated by the originating network via STIR.

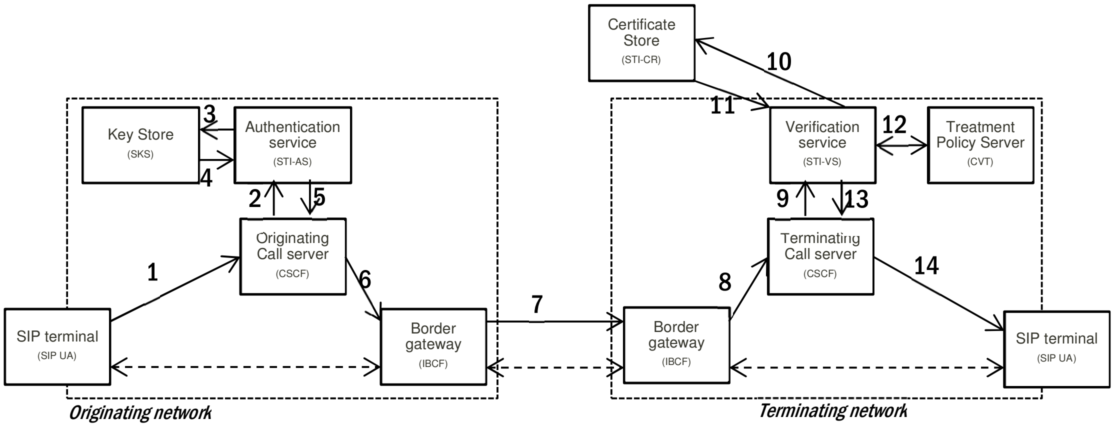

Figure 6.2.a: STIR call flow

During the interim stage of STIR, the call flow will be as follows:

1. The originating SIP terminal, which is registered and authenticated to the Originating Call Server, creates a SIP INVITE with a telephone number identity in the From: header that, according to customer configuration, may be intended to be used as the PN for display purposes.

2. The Originating Call Server populates the P-Asserted-Identity header field asserting the Network Number CLI of the originating SIP Terminal's network termination point. It also populates the From: header field representing the Presentation Number CLI to be displayed for the call:

    a. In the case of a customer utilising a Type 2, 3, 4 or 5 Presentation Number CLI received from the SIP terminal in the From: header field, then this is used to populate the outgoing From: header field.

    b. In all other cases, the Presentation Number will be a statically configured value for that SIP terminal.

3. The Originating Call Server then initiates an originating trigger to the Authentication Service for the INVITE.

4. The Authentication Service first determines the legitimacy of the PN CLI being used in the INVITE (during the Start-up and Interim Stages, it will be an individual operator matter what the criteria for this is).

    - If acceptable, the Authentication Service then securely requests its private key for the Network Number from the Key Store.

    - If unacceptable then the Authentication Service could take various actions: it could

        i. fail the call,

        ii. pass the call with an unsigned Network Number, or

iii. insert a suitable Presentation Number (which may be a copy of the Network Number) and then securely requests its private key for the Network Number from the Key Store.

5. The Key Store provides the private key in the response, and the Authentication Service signs the Identity header field in the INVITE as specified in RFC8224 using the Network Number CLI in the P-Asserted-Identity header field. Where the contents of the From: header field are provided by the originating network or validated against a whitelist by the Authentication Service, the signing will be marked as “Full attestation”; otherwise, it will be marked as “Partial attestation”.
6. The Authentication Service passes the INVITE to the Originating Call Server.
7. The Originating Call Server, through standard resolution, routes the call to the egress Border Gateway.
8. The INVITE is routed over the NNI through the standard inter-domain routing configuration.
9. The Terminating Network’s ingress Border Gateway receives the INVITE over the NNI.
10. The Terminating Call Server initiates a terminating trigger to the Verification Service for the INVITE.
11. The Terminating Verification Service uses the “info” parameter information in the Identity header field as specified in RFC8224 to determine the Certificate Store Uniform Resource Identifier (URI) and makes an HTTPS request to the Certificate Store.
12. The Verification Service validates the certificate then extracts the public key. It constructs the RFC8224 format information and uses the public key within this to verify the signature in the Identity header field, which validates that the Network Number CLI used is authentic. In the Start-up phase, the Verification Service must rely on the trustworthiness of the originating network in order to assess whether the Network Number CLI is reasonable for that network. In the Interim phase, once a common database of numbers is available, this will be used to determine whether the originating network has the right to use the Network Number CLI. The authenticity of the Network Number CLI is used to assess the likely authenticity of the Presentation Number CLI information in the From: header field.
13. The Treatment Policy Server is an optional function that can be invoked to perform call spam analytics or other mitigation techniques and return a response of how the call should be treated.
14. Depending on the result of the verification, the Verification Service determines whether the call is to be completed, and if so with any appropriate indicator, and the INVITE is passed back to the Terminating Call Server which continues to set up the call to the destination SIP Terminal.
15. The destination SIP Terminal receives the INVITE and normal SIP processing of the call continues.

The final stages of STIR implementation would change the call flow according to which of the variants set out in Section 5 is subsequently adopted. Annex B sets out the changes that would be required to the call flow.

6.3 Terminating Network actions

A terminating network which supports STIR will need to define what action to take when receiving a call, dependent on the level of trust in the CLI established by the STIR implementation. Where technically feasible, CPs may wish to offer their customers an individual choice on how they would like calls of each category to be handled, as shown in Table 6.3.a.

Table 6.3.a: Example options for Terminating Network

Level of trust	**Options for Terminating Network**
None	Route call with no displayed PN Route call, with an advice whisper on answer Route call with a visual advice of un-validated PN (for IP phones with such capability) Send call to voicemail box Block call (based on user opt-in to blocking) Send call through a screening service before routing Block call (based on CP choice dependent on upstream carrier)
Partial	Route call with or without displayed PN Route call, with a cautionary advice whisper on answer Route call with a visual advice of low PN reliability score or ‘trust marking’ (for IP phones with such capability) Send call to voicemail box Block call (based on user opt-in to blocking) Send call through a screening service before routing Block call (based on CP trust level of originating CP)
Full	Route call with displayed PN Route call with visual advice of high PN reliability score or ‘trust marking’ (for IP phones with such capability) Route call with ‘trusted caller’ whisper upon answer.

The conditions leading to a given level of trust would evolve as STIR is implemented, as shown in Table 6.3.b. In determining this approach, the assumption is that conditions resulting in no trust would be those where regulation is being breached, and those with partial trust would be those where the terminating network can’t verify that the CLI is absolutely reliable, but has some evidence to suggest it may be.

Table 6.3.b: Options for Terminating Network: conditions for trusting CLI

Stage/Phase - Note 1	Conditions resulting in no trust	Conditions resulting in partial trust – Note 2	Conditions resulting in full trust – Note 2
Start-up stage (phase 2a)	N/A	a. No signing of NN CLI, or b. NN CLI is signed with gateway attestation, or c. NN CLI is signed with partial attestation but originator is untrusted d. NN CLI is signed with full attestation but originator is untrusted	a. NN CLI is signed with full attestation and originator is trusted b. NN CLI is signed with partial attestation and originator is trusted
Interim Stage (phase 2b)	N/A	a. No signing of NN CLI, or b. NN CLI is signed with gateway attestation, or c. NN CLI is signed with partial attestation and originator is untrusted	a. NN CLI is signed with full attestation b. NN CLI is signed with partial attestation and originator is trusted
Interim Stage (phase 4) – Note 3	No signing of NN CLI	a. NN CLI is signed with gateway attestation, or b. NN CLI is signed with partial attestation and originator is untrusted	a. NN CLI is signed with full attestation b. NN CLI is signed with partial attestation and originator is trusted
Final Stage (phase 6I/6II) – Note 4	a. No signing of NN CLI, or b. NN CLI is signed with partial attestation (Note 5)	NN CLI is signed with gateway attestation	NN CLI is signed with full attestation
Final Stage (phase 6III) – Note 6	a. No signing of either PN or NN CLI b. NN CLI signed with partial attestation (Note 5)	NN CLI is signed with gateway attestation	a. NN CLI is signed with full attestation, or b. PN CLI is signed

Notes:

1. The stage/phases used in this table have been chosen as they represent the point at which the terminating network's trust in originating networks changes (see Annex A for further information on these phases)
2. In all cases, it is an individual terminating network operator matter to set suitable criteria for assessing the trustworthiness of originating networks.
3. By this time the expectation is that all Network Number CLIs would be signed.

4. Under these Final Stage Variants, the originating network should never be signing with partial attestation, and instead should be signing with full attestation, after having either verified the supplied PN against a whitelist (variant 6I) or checking a customer-signed PN (variant 6II).
5. The terminating network could optionally treat this case as having partial trust according to the level of trust they have in the originating CP.
6. Under Final Stage Variant 6III, any customer supplied PN should be signed by that customer or treated as untrusted.

7 Distribution of credentials

The architecture and call flow in Section 6 set out that an Authentication Service in the originating network digitally signs the CLI, and a Verification Service in the terminating network checks that signature. This means that there is a need for arrangements for the Authentication Service to have a private key to sign the CLI, and the Verification Service to have an associated public key in order to check that signature. In the context of Figure 6.2.a, there is a need for an approach to be agreed for how the private and public keys will be populated into the Key Store and Certificate Store, and also for agreements of where these functions will practicably reside.

NICC has examined a series of options for credential distribution and storage, which are analysed in Annex C. In the earlier version of this report, it was concluded that the most promising option for UK implementation was Approach 5, set out in Figure 7.a below, which foresaw certificates being associated with groups of numbers, and a check of rights-to-use of numbers being carried out prior to certificates being issued. However, having reviewed the US implementation, and considered the complexity involved in launching STIR, NICC now recommends Approach 2, set out in Figure 7.b below. This is because:

- Whereas Approach 5 means that a numbering database is required before certificates can even be issued to facilitate the signing of CLIs, Approach 2 allows a start-up stage without a database, albeit with reduced usefulness.
- When the numbering database is made available, terminating operators can choose when to make use of it to enhance their Verification Service, aligned with when they might choose to make use of its contents as an originating network.

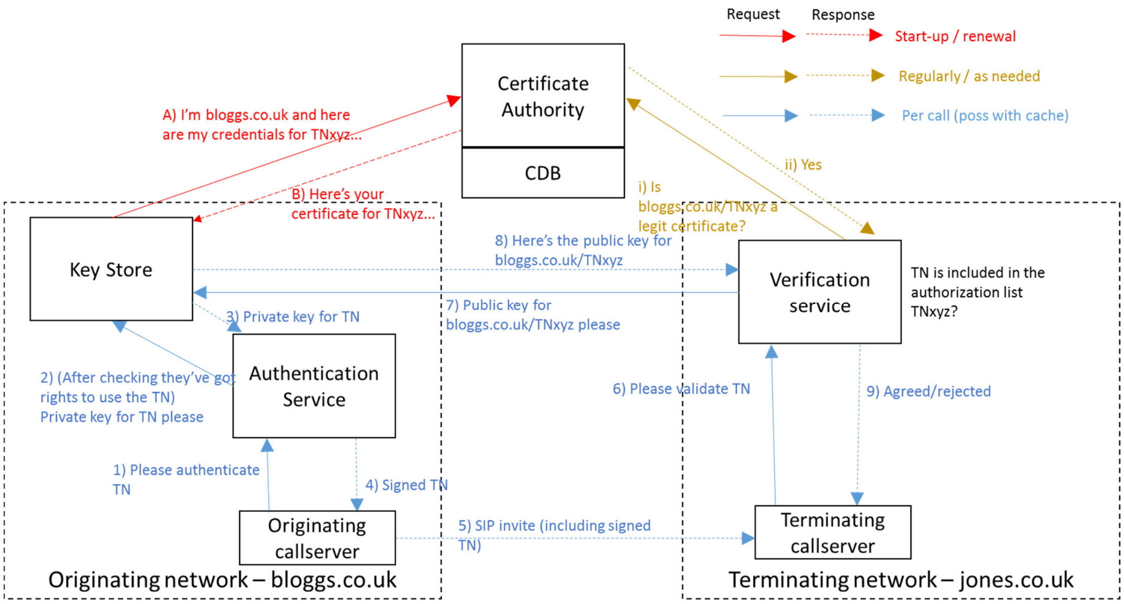
Request
Response
Start-up / renewal
Regularly / as needed
Per call (poss with cache)
Certificate Authority
CDB
A) I'm bloggs.co.uk and here are my credentials for TNxyz...
B) Here's your certificate for TNxyz...
i) Is bloggs.co.uk/TNxyz a legit certificate?
ii) Yes
Key Store
8) Here's the public key for bloggs.co.uk/TNxyz
Verification service
TN is included in the authorization list TNxyz?
3) Private key for TN
7) Public key for bloggs.co.uk/TNxyz please
2) (After checking they've got rights to use the TN) Private key for TN please
Authentication Service
6) Please validate TN
9) Agreed/rejected
1) Please authenticate TN
4) Signed TN
Originating callserver
5) SIP invite (including signed TN)
Terminating callserver
Originating network – bloggs.co.uk
Terminating network – jones.co.uk

Figure 7.a: Earlier preferred approach (Approach 5) – now not the recommended option

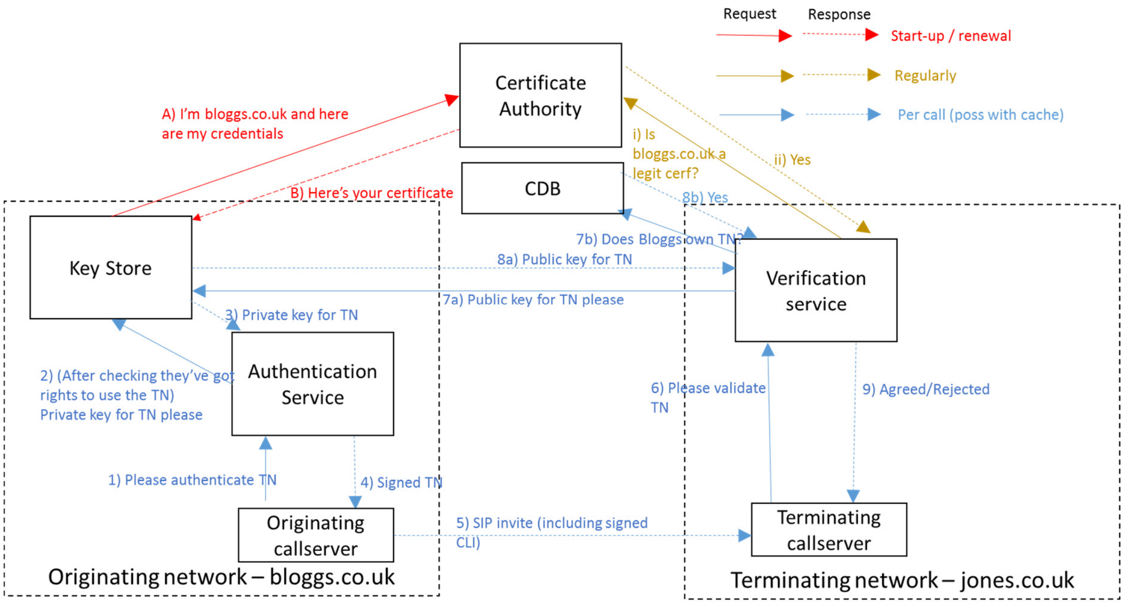
Request
Response
Start-up / renewal
Regularly
Per call (poss with cache)
Certificate Authority
A) I'm bloggs.co.uk and here are my credentials
i) Is bloggs.co.uk a legit cerf?
ii) Yes
CDB
B) Here's your certificate
8b) Yes
7b) Does Bloggs own TN?
8a) Public key for TN
Key Store
Verification service
7a) Public key for TN please
3) Private key for TN
2) (After checking they've got rights to use the TN) Private key for TN please
Authentication Service
6) Please validate TN
9) Agreed/Rejected
1) Please authenticate TN
4) Signed TN
Originating callserver
5) SIP invite (including signed CLI)
Terminating callserver
Originating network – bloggs.co.uk
Terminating network – jones.co.uk

Figure 7.b: Recommended STIR credentials treatment for UK

Under this preferred approach, there would be a central Certificate Authority for the UK numbering plan, which would distribute the certificates that are then used to generate public and private keys for usage in STIR.

The Certificate Authority would issue the certificates to operators solely based upon them being a *bona fide* operator: in this context NICC recommends that the Certificate Authority would be appointed by and have a strong relationship with Ofcom. The Originating Network would use this certificate to create keys to sign the CLI. The Verification Service would need to carry out two checks, namely:

1. Whether the certificate associated with signing the CLI is valid – this is shown in brown in Figure 7.b and in practical terms would be implicit in the certificate signing infrastructure rather than being an overt information exchange, and

2. Whether the originating network has rights to use/sign that number; this requires a common numbering database and although Figure 7.b depicts this check as an external query from the terminating network, in practical terms it would probably accomplish this by examining its local copy of the database. As set out in Section Five, in the Start-up Stage where there is no numbering database, this check would not be possible and the terminating network would need to take a judgement based upon their trust of the originator.

8 International calls

8.1 Inbound international calls

Whilst the above approach set out in Section 7 is suitable for nationally originated calls, it is less so for internationally originated. The treatment of such calls will depend upon whether the call signalling received by the international gateway contains a P-A-ID header field, and if so whether it has been signed by a preceding network.

It should be noted that on the whole, received P-A-ID header fields will contain non-UK numbers, but there will be exceptions, for example roaming mobile numbers and potentially calls that have been subject to least-cost routeing hence exited and re-entering the UK. However, calls from fixed lines originated outside of the UK should not contain P-A-ID header fields containing a UK CLI, as a UK Network Number should only represent a UK Network Termination Point; this differs from an internationally-originated From: (Presentation Number) parameter, which could legitimately contain a UK number.

Inbound call contains a P-A-ID header field containing an invalid CLI
As set out in ND1016 [3], the call would be blocked.

Inbound call contains a P-A-ID header field containing a valid CLI
The following treatments would apply:

> *P-A-ID header field is already signed*
> In this situation it is recommended that as a long term solution the gateway node would pass the signed CLI through transparently, and the terminating network would then seek to verify this information. If the CLI concerned is a non-UK number, the terminating network's verification would need to rely upon an overseas Certificate Authority. This will require international agreement to allow terminating network Verification Services to download and install the root certificate for approved Certificate Authorities.
>
> It should be noted that the P-A-ID header field may contain a number that (pre-STIR) would have led to the call being blocked based on the international gateway's assessment of the reliability of that CLI; under this long-term model the call would be passed with that P-A-ID header field left intact for the terminating network to make this judgement based upon the signing. This could be considered a disadvantage compared to the current approach, but it leaves the decision in the hands of the called party and/or terminating network, rather than a third party.
>
> However, in the short-medium term, passing the call through with signed information for a terminating network to interpret could be a retrograde step. If the terminating network, or any UK transit network in the call-path, is not using IP technology, the STIR information would be lost. This means that a call that could have been blocked at the international

gateway as containing an untrustworthy CLI would be routed to the terminating customer. Therefore, until the international gateway can be sure that downstream networks are capable of processing the received STIR information, the call blocking procedures set out in ND1016 [3] rule NC1 would need to prevail. Note, however, that the international gateway would still pass through the STIR information on calls that are not subject to blocking, in order that any terminating networks that do support STIR can process this.

P-A-ID header field is unsigned but the gateway considers the CLI reliable
In this situation the gateway network would have received a P-A-ID header field that it had no reason to mistrust. The gateway node would sign the number as gateway certified and the terminating network then make its judgement based upon the attestation being set to gateway (see Section 6.3). The check from the Verification Service to the Certificate Authority depicted in Figure 7.b would merely verify that the gateway network is known by the Certificate Authority (i.e. there would be no check of rights of use specific numbers).

P-A-ID header field is unsigned and the gateway does not consider the CLI reliable
ND1016 [3] rule NC1 dictates that in this situation the gateway network would block the call. This behaviour should continue for the short-medium term (i.e. the gateway network makes a judgement as to the trustworthiness of the CLI), for the reasons set out above with respect to calls received with a signed P-A-ID Header field. In the long term, consideration would need to be given whether to continue this approach, or instead the gateway network would insert a valid P-A-ID header field containing a number from the 0897 number range indicating where the call entered the UK, and sign that CLI with gateway attestation, showing it had a lesser level of trust in the call than those signed with full attestation. The logic of such a change would be that all blocking would be focussed on the terminating network serving the customer, but this would be at the expense of the loss of any information the gateway provider would have of the upstream international network.

Inbound call contains no P-A-ID header field
ND1016 [3] rule NC1 dictates that in this situation the gateway network would be inserting a valid P-A-ID header field containing a number from the 0897 number range, indicating where the call entered the UK. Were STIR to be adopted, the gateway network would sign that inserted number with gateway attestation.

8.2 Outbound international calls

Other than where ND1016 rule NC2 is invoked to remove a CLI or where the destination network requests that no STIR signalling be received, it is recommended that any STIR signalling is passed transparently. If an international terminating network then wishes to validate the CLI, this would imply that they would need to establish a relationship with the UK Certificate Authority.

If ND1016 rule NC2 is invoked and CLI information is removed, then the international gateway node should also remove the signed CLI information.

9 Benefits and gap analysis of STIR

This section assesses how well an implementation of STIR as set out in this document would achieve the goal of assuring that the caller has the right to use the CLI presented to the called customer. Table 9.a sets out what the implementation would achieve at each stage for each of the CLI types that were described in Section 4; text in amber shows where the interim stage does not provide full assurance of the displayed CLI, whereas text in red shows where even the final stage does not meet this goal.

Table 9.a: Efficacy of a UK STIR implementation

CLI Scenario	Start-up Stage (no database)		Interim Stage		Final Stage Variant I (Originating network whitelists acceptable PN CLIs)		Final Stage Variant II (Originating network validates customer-signed PNs)		Final Stage Variant III (Terminating network validates customer-signed PNs)		Comments
	Treatment	What does this give	Treatment	What does this give	Treatment	What does this give	Treatment	What does this give	Treatment	What does this give	
1. Residential line (no PN)	NN signed by Originating Network with full attestation	Confirmation that the CLI is reliable/authentic (if the terminating network knows and trusts originating network)	NN signed by Originating Network with full attestation	Confirmation that the CLI is reliable/authentic	Same as Interim stages						Assuming SIP, then From header field is populated with the NN (i.e. is the same as the P-A-ID header field)
2. Business line (no PN)	NN signed by Originating Network with full attestation	Confirmation that the CLI is reliable/authentic (if the terminating network knows and trusts originating network)	NN signed by Originating Network with full attestation	Confirmation that the CLI is reliable/authentic	Same as Interim stages						Assuming SIP, then From header field is populated with the NN (i.e. is the same as the P-A-ID header field)
3. Business line, static PN (Type 1)	NN signed by Originating Network with full attestation	Confirmation that the PN is reliable/authentic (if the terminating network knows and trusts originating network)	NN signed by Originating Network with full attestation	Confirmation that the PN is reliable/authentic	Same as Interim stages						

CLI Scenario	Start-up Stage (no database)		Interim Stage		Final Stage Variant I (Originating network whitelists acceptable PN CLIs)		Final Stage Variant II (Originating network validates customer-signed PNs)		Final Stage Variant III (Terminating network validates customer-signed PNs)		Comments
	Treatment	What does this give	Treatment	What does this give	Treatment	What does this give	Treatment	What does this give	Treatment	What does this give	
4. Business line where the originating network has verified the number received from the customer (Type 2)	NN signed by Originating Network with full attestation	Confirmation that the PN is reliable/authentic (if the terminating network knows and trusts originating network)	NN signed by Originating Network with full attestation	Confirmation that the PN is reliable/authentic	Same as Interim stages						Assumes quality of checking in originating network is sufficient
5. Business line, enterprise network with multiple sites/extensions (Type 3)	NN signed by Originating Network with partial attestation	Confirmation of the originating CP that allowed the PN into the public network	NN signed by Originating Network with partial attestation	Confirmation of the originating CP that allowed the PN into the public network	Originating Network validates number against whitelist, then signs NN with full attestation	Confirmation that the PN is reliable/authentic	PN signed by the enterprise with full attestation. Originating Network checks this and then signs NN with full attestation	Confirmation that the PN is reliable/authentic	PN signed by the enterprise with full attestation. Terminating Network checks this	Confirmation that the PN is reliable/authentic	Interim stage confirm identity of originating network but not necessarily validity of the PN. For variant I enterprise has to inform originating network of any additions
6. Business line, enterprise network allowing break in-break out (Type 4)	NN signed by Originating Network with partial attestation	Confirmation of the originating CP that allowed the PN into the public network	NN signed by Originating Network with partial attestation	Confirmation of the originating CP that allowed the PN into the public network	NN signed by Originating Network with partial attestation	Confirmation of the originating CP that allowed the PN into the public network	NN signed by Originating Network with partial attestation	Confirmation of the originating CP that allowed the PN into the public network	NN signed by Originating Network with partial attestation	Confirmation of the originating CP that allowed the PN into the public network	Cannot reliably sign the PN CLI as it doesn't belong to the enterprise
7. Business line, call-centre passing different PN CLIs according to campaign or client (Type 5)	NN signed by Originating Network with partial attestation	Confirmation of the originating CP that allowed the PN into the public network	NN signed by Originating Network with partial attestation	Confirmation of the originating CP that allowed the PN into the public network	Originating Network validates number against whitelist, then signs NN with full attestation	Confirmation that the PN is reliable/authentic	PN signed by the enterprise with full attestation. Originating Network checks this and then signs NN with full attestation	Confirmation that the PN is reliable/authentic	PN signed by the enterprise with full attestation. Terminating Network checks this	Confirmation that the PN is reliable/authentic	Interim stages confirm identity of originating network but not necessarily validity of the PN. For variants II and III, call-centre must sign using credentials of their client. For variant I, call-centre must inform originating network of any additions

CLI Scenario	Start-up Stage (no database)		Interim Stage		Final Stage Variant I (Originating network whitelists acceptable PN CLIs)		Final Stage Variant II (Originating network validates customer-signed PNs)		Final Stage Variant III (Terminating network validates customer-signed PNs)		Comments
	Treatment	**What does this give**	**Treatment**	**What does this give**	**Treatment**	**What does this give**	**Treatment**	**What does this give**	**Treatment**	**What does this give**	
8. Call Diversion within network	Per relevant row above	Per relevant row above	Per relevant row above	Per relevant row above	Per relevant row above	Per relevant row above	Per relevant row above	Per relevant row above	Per relevant row above	Per relevant row above	STIR signing rippled through by diverting network
9. Call Diversion by customer equipment	NN signed by the network hosting the customer diverting the call, with partial attestation	Confirmation of the network hosting the customer diverting the call, that allowed the PN into the public network	NN signed by the network hosting the customer diverting the call with partial attestation	Confirmation of the network hosting the customer diverting the call, that allowed the PN into the public network	NN signed by the network hosting the customer diverting the call with partial attestation	Confirmation of the network hosting the customer diverting the call, that allowed the PN into the public network	NN signed by the network hosting the customer diverting the call with partial attestation	Confirmation of the originating CP that allowed the PN into the public network	NN signed by the network hosting the customer diverting the call with partial attestation	Confirmation of the originating CP that allowed the PN into the public network	Cannot reliably sign or verify the PN CLI as it doesn't belong to the enterprise
10. UK mobile on home network	NN signed by Originating Network with full attestation	Confirmation that the CLI is reliable/authentic (if the terminating network knows and trusts originating network)	NN signed by Originating Network with full attestation	Confirmation that the CLI is reliable/authentic	Same as Interim stages						Assuming SIP, then PN (From) is populated with the NN (P-A-ID header field)
11. UK mobile roaming overseas (home network routeing)	NN signed by Home Network with full attestation	Confirmation that the CLI is reliable/authentic (if the terminating network knows and trusts home mobile network)	NN signed by Home Network with full attestation	Confirmation that the CLI is reliable/authentic	Same as Interim stages						Call routes via home network which behaves as originating network for this purpose – NB this is the default for VoLTE
12. UK mobile roaming overseas (direct routeing)	NN signed by inbound international network with gateway attestation	Confirmation of which international gateway provider accepted call into the UK	NN signed by inbound international network with gateway attestation	Confirmation of which international gateway provider accepted call into the UK	NN signed by inbound international network with gateway attestation	Confirmation of which international gateway provider accepted call into the UK	NN signed by inbound international network with gateway attestation	Confirmation of which international gateway provider accepted call into the UK	NN signed by inbound international network with gateway attestation	Confirmation of which international gateway provider accepted call into the UK	NB this isn't the standard model for VoLTE. If it became so then for further study – could the visited network sign using credentials passed by home network?

CLI Scenario	**Start-up Stage (no database)**		**Interim Stage**		**Final Stage Variant I (Originating network whitelists acceptable PN CLIs)**		**Final Stage Variant II (Originating network validates customer-signed PNs)**		**Final Stage Variant III (Terminating network validates customer-signed PNs)**		**Comments**
	Treatment	**What does this give**	**Treatment**	**What does this give**	**Treatment**	**What does this give**	**Treatment**	**What does this give**	**Treatment**	**What does this give**	
13. Foreign mobile roaming in UK			?				?		?		It is unclear what would happen in this call case; potentially this has to be gateway attestation as visited network has no rights to the NN
14. Overseas call-centre wishing to display UK number: connected to UK network (i.e. long lined)	NN signed by Originating Network with partial attestation	Confirmation of the originating CP that allowed the PN into the public network	NN signed by Originating Network with partial attestation	Confirmation of the originating CP that allowed the PN into the public network	Originating Network validates number against whitelist, then signs NN with full attestation	Confirmation that the PN is reliable/authentic	PN signed by the enterprise with full attestation. Originating Network checks this and then signs NN with full attestation	Confirmation that the PN is reliable/authentic	PN signed by the enterprise with full attestation. Terminating Network checks this	Confirmation that the PN is reliable/authentic	Interim stage confirms identity of originating network but not necessarily validity of the PN. For variant I enterprise has to inform originating network of any additions
15. Overseas call-centre wishing to display UK number: connected to overseas network	NN signed by inbound international network with gateway attestation	Confirmation of which international gateway provider accepted call into the UK	NN signed by inbound international network with gateway attestation	Confirmation of which international gateway provider accepted call into the UK	Originating Network validates number against whitelist, then signs NN with full attestation	Confirmation that the PN is reliable/authentic, but only if the overseas originating network is trusted	PN signed by the enterprise with full attestation. Originating Network checks this and then signs NN with full attestation	Confirmation that the PN is reliable/authentic, but only if the overseas originating network is trusted	PN signed by the enterprise with full attestation. Terminating Network checks this	Confirmation that the PN is reliable/authentic, but only if overseas networks pass STIR information	Variants I and II relies on a) trusting signed non-UK NN CLI and having infrastructure to verify it and b) trusting overseas network on full vs partial attestation Variant III relies on the availability of a global certification infrastructure for enterprises.

As can be seen from Table 9.a, the start-stage works on the basis of a “circle of trust” between a subset of originating and terminating networks – this could be on a reciprocal agreement basis, but could present issues of unequal power (i.e. it is in a small operator’s interests to trust signing by a large operator, but not necessarily the other way around). Ofcom should consider whether it should adopt procedures for where disputes arise as to trustworthiness. For calls that are originated in the UK, the Startup Stage allows the originating network admitting the Presentation Number into the public network to be identified, which will allow terminating networks to place pressure on originators should there be evidence of illegal CLI spoofing.

The interim stage goes a long way to restoring the integrity of CLIs by building in a check of whether the originating network had the rights to at least use the Network Number. However, what it does not provide is real-time validation that the Presentation Number is legitimate. The interim stage also relies on a degree of trust that the originating network will behave correctly and only assert that a CLI is signed with full attestation if it can be sure that the Presentation Number is valid; if originating networks sign Presentation Numbers received from enterprise customers as being fully attested rather than with the correct partial attestation, this will compromise the efficacy of STIR.

At this time, it is not clear whether this capability gap will be sufficient to justify a move to one or more of the Final Stage Variants. If it is, then the Final Stage Variants will allow most UK customer-supplied Presentation Numbers to be validated, but once again there must be a degree of trust that originating networks will correctly do this. There would, however, be a shortfall for Type 4 CLIs, i.e. private network break-in/out, and calls diverted by a private network. For these call cases, in principle the Presentation Number could contain any number and therefore:

- For Final Stage Variant I, the originating network is unable to establish a whitelist of CLIs because this list would contain every possible number.
- For Final Stage Variants II and III, the customer would be unable to sign the Presentation Number as it is not their number.

As such, these call cases would go unsigned and if they are the only national call-case not signed, would likely be rejected as suspicious by terminating networks.

A further call case that could be problematic is that of roaming mobile terminals where direct routeing is invoked (i.e. the calls route directly from the visited network). In this scenario, any calls to the UK could only by signed as “gateway attested” at the inbound international node, unless some mechanism can be found to pass signing information to the visited network and for them to carry out the signing. This said, the principal model for VoLTE roaming appears to be that calls are routed via the home network, so the exposure may be limited.

Finally, there is a significant gap where calls are originated from overseas call-centres using a UK number. These calls could be authenticated only if the originating network is brought within the “trust circle” of a UK STIR implementation. Arguably, however, it is this call scenario which is the dominant source of unsolicited marketing calls, so being able to merely differentiate such calls as non-STIR validated could be considered an advantage. Further, if the call-centre owner in question is concerned about calls being rejected due to not being STIR validated, there are avenues open to them such as on-shoring the call-centre, or long-lining it for egress into a UK public network.

10 Implications for Network Protocols

10.1 UK SIP support for STIR PASSporT

The STIR mechanism (as specified by RFC8224) uses a SIP Identity header field containing a STIR PASSporT (as specified in RFC8225 [7]), which is a token format that provides (among other things) a signature over the Date header field of SIP requests and parts of the To and From header fields. SHAKEN extends the PASSporT definition to include attestation claims (as specified by ATIS SHAKEN/draft-wendt-stir-passport-shaken-00).

In order to support STIR in UK SIP networks, it is essential that the SIP Identity header field can be passed end to end between the participating originating and terminating SIP entities. This is primarily a requirement for ND1035 SIP NNI [4]; also for a SIP UNI such as ND1034 [8] if/when signing/validation functions are extended to customer/user entities.

10.2 Compact vs full form

Specific claims within the STIR PASSporT relate to other elements of the SIP request:

- "orig" (Origination Identity) – derived from the From header field (for basic STIR); or from the P-Asserted-Identity header field (if present) in the case of the SHAKEN extension of STIR.
- "dest" (Destination Identity) – derived from the To header field.
- "iat" (Issued at) – derived from the Date header field.

For example:

```
{ "orig":{"tn":"12155551212"},
  "dest":{"tn":"12155551213"},
  "iat":1443208345 }
```

The PASSporT structure is defined to have two alternative formats: full form and compact form. The full form includes the complete content of the PASSporT, whereas the compact form contains only the signature part. The compact form reduces message size; however, when using this form it is important that the relevant header field content (from which the "dest" and "iat" claims are derived) is passed end-to-end without modification in order for the signature to be successfully validated by the called user's Verification Service. There are valid use cases in UK (and other) networks which can result in the To and/or Date header field being modified in transit. Therefore it is proposed that UK implementation of STIR would use the full form PASSporT.

10.3 PASSporT URI

According to RFC8224, the "info" parameter of the PASSporT contains a URI which dereferences to a resource that contains the public key components of the credential used by the authentication service to sign a request. The URI must conform to one of three URI schemes (according to draft-ietf-stir-certificates-14): the CID URI, the SIP URI, and the HTTPS URI.

For UK implementation of STIR, the "info" parameter must contain a HTTPS URI.

10.4 Other considerations

The following observations may be relevant to future UK applications of STIR.

- STIR only protects the identity part of the From/P-A-ID header fields (for NICC purposes, a telephone number). Deleting or adding parameters (e.g. CPC) will not be detectable from the signature.
- Similarly, the STIR PASSporT does not explicitly indicate whether the signed identity is that of the P-A-ID or the From header field. So, for example:
  - If the P-A-ID header field is present, then the content of the From is not protected against modification at an intermediate node.
  - There is no protection against insertion of a P-A-ID header field identical to the From in the case where a P-A-ID header field was not originally present (and hence the "orig" claim reflects the From content). (For UK networks compliant with ND1016 this should not be a problem, as the P-A-ID header field should be inserted (and suitably signed) by the originating UK CP; however it may be consideration for interworking with non-UK networks.)
- The STIR Passport payload must contain the JSON key "iat" – (Issued At claim). The “iat” key should be computed from the original SIP Date header field and is encoded using UNIX time format as per RFC 7519. RFC8224 recommends that a local freshness validity policy of 60 seconds from computation should be adopted in relation to “iat”. This is to mitigate replay attacks

11 Implementing the STIR Functions

The STIR functions set out in the previous sections can be separated into those that are within individual communication provider domains, and those which will need to be operated by a (probably independent) third party for the benefit of all communications providers. NICC is not in a position to speculate on quantitative costs of implementing these functions, however this section provides an overview of the likely complexity of them.

11.1 Communication Provider Functions

11.1.1 Call Server and Border Gateways

The main change to Call Server functionality (or, as relevant according to the individual network implementation, Border Gateway functionality) will be to accommodate the population and handling of the additional fields representing the signed CLI within SIP signalling, and (assuming this isn’t already present) an additional query to an application server. The cost of this will very much depend upon whether STIR technology is adopted internationally (ideally via 3GPP), hence making this standard functionality, or whether the UK is largely going alone in adopting. Should the latter be the case, then the costs may be prohibitive.

Changes will need to be made to Border Gateway functions to pass both STIR signalling, and all parameters that are signed by STIR signalling, transparently between networks. This is a material change to existing call logic, but adoption of SHAKEN in America may ease implementation.

11.1.2 Authentication & Verification Services

It is envisaged that both the authentication and verification services will be accommodated via application server functionality in networks. Once again, the question of cost will significantly depend upon international adoption. Even with international adoption, however, the application servers concerned will need to be involved in every originating/terminating call setup, so will need to be of substantial capacity.

11.1.3 Key Store

The key store functionality is somewhat of a bridge between the telephony and internet worlds – to date the equivalent functionality has operated on internet domain & sub-domain names rather than numbers.

Ultimately, if the certificate challenges are addressed, then it is likely that the key store will be largely standard functionality (if STIR is adopted widely internationally), but the costs associated with the functionality are unclear at this time.

11.2 Central Functions – Certificate Authority

The Certificate Authority functionality will need to be implemented by one or more agencies that will need to be authorised by Ofcom or UK communication providers acting collectively. Certificate Authorities are widely implemented for internet domain names, but STIR functionality, particularly as envisaged by the recommended option in this report, would need a definitive validation that certificates are being issued only to specific network entities authorised by Ofcom.

The Certificate Authority as envisaged in Section 7 of this report would not need to be queried in real-time, and as such does not need 99.999% resilience.

12. Alternatives to STIR

Section 9 identified that although adoption of STIR technology has the scope to improve the reliability of CLI, there are still substantial gaps, particularly until such a time that signing of CLIs is implemented internationally, with interworking between the national trust domains. In particular, for inbound international calls, a UK-only implementation of STIR will merely serve to indicate which international gateway brought the call into the UK. Moreover, as set out in Section 10 and 11, implementation of STIR has significant implementation costs. Whilst not meaning that STIR shouldn’t be implemented, it does raise the question of whether there are alternative lower cost approaches that could bring the advantages of STIR until such a time that there was wide-scale international adoption.

The STIR approach has been designed to allow the identity of the originating network to be conveyed, while countering two specific threats:

1. That an originator passes themselves off as someone else in order to generate nuisance calls, and
2. That a man-in-the-middle changes the identity of the signalled originating network for their own ends.

The first threat is addressed by calls being signed with a certificate from a recognised authority, and the second by the information being signed hence not vulnerable to being changed.

A lower cost short-medium term solution for the UK may be to adopt an approach that mitigates these threats via process means. It may be possible to adopt a parameter in SIP signalling – ideally without needing to devise a new header – that identifies the network that admitted a call into the

“UK trust domain”. This would be the originating network in the case of a nationally-originated call, or the international gateway in the case of an inbound international call (it could also be possible to block inbound international calls with a UK CLI).

Without the sophistication of STIR technology, this parameter would of course be vulnerable to the threats above. However, measures could be taken to reduce the risk: in the case of an originator passing themselves off as a third party, then transit networks (and terminating networks where there was no transit) could police whether the parameter value was appropriate for the interconnection concerned. So if a route from originating network A contained calls with the parameter populated as originating network B being the source of the calls, then action could be taken. This policing and action could be on a per-call basis, but more likely would be either as an audit function, or alternatively as a result of complaints from terminating networks.

Similarly, although the UK has not particularly experienced man-in-the-middle attacks as set out above, they could be recognised and remedied should a terminating network approach an originating network (according to the value of the originating network parameter) and find that they had not in fact originated the call, hence the parameter had been tampered with.

This approach would also allow terminating networks to establish logic to handle calls according to the trustworthiness of the originating network, which could be modified to be driven by STIR signalling as and when that is adopted. It is also compatible with the usage of a CDB to check whether an originating network should be using a particular P-A-ID, independent of whether STIR is adopted. In advance of STIR, it would readily identify that the entry point into the UK trust domain is an international gateway, and allow terminating networks to apply their own blocking policy based on that knowledge.
NICC will be studying whether such a parameter could be implemented as a lower cost alternative to STIR in the short term.

13 Conclusions and Next Steps

This document has set out what would be achieved by a staged implementation of STIR technologies in UK networks. The Start-up Stage would in principle provide surety of which originating CP admitted a call into the public network, and the Interim Stage would add a check that the Network Number CLI is one which the originator is permitted to use. It would not, however, provide ultimate confirmation that the Presentation Number CLI is one which the originator is permitted to use (in the case of Type 3/4/5 Presentation Number CLIs), and it is these call cases which are at the root of much nuisance calling. Only a move to one of the Final Stage solutions would extend confidence to Presentation Number CLIs.

It is possible that being able to trace calls to a given originator could, however, inject such a level of transparency of the source of calls that the prevalence of UK-originated nuisance calls could be reduced meaning there is insufficient justification to move to the Final Stage for UK-originated calls.

Notwithstanding this, there are certain call cases which will not be addressed by STIR, either in an Interim or Final Stage solution. Calls from overseas call-centres, which represent a large proportion of nuisance calls, are not well-addressed. Also scenarios such as calls diverted by enterprise networks could not have properly signed CLI information. Arguably if STIR is implemented, a non-signed CLI could be an indicator that the caller is suspect, but it would be difficult to distinguish between calls with unsigned CLIs because they originate outside the trust domain (so possibly nuisance), and those which are a call-case which STIR is ill-equipped to serve (diverted calls).

In summary, implementing STIR technology in UK networks could be a large step forward, but it is certainly not a panacea for resolving loss of trust in CLI, let alone for preventing nuisance calls. Any technical solution would need to be accompanied by regulatory action both to mandate implementation of STIR and against the perpetrators of nuisance calls.

Any implementation of STIR would necessarily be a long-term task in the UK. It depends upon availability of end-to-end SIP, and this needs to be universal in order to allow terminating networks to take any action based upon the presence of a signed CLI (otherwise, it wouldn't be possible to distinguish between lack of signing because the CLI can't be trusted, and lack of signing because the call path isn't end-end SIP). Although many networks are now SIP-based, for some large communications providers this is a number of years away. As such, although NICC considers that signing by IP originating networks could commence in the short term (e.g. 2021), it will be many years before the lack of signing could be taken to be an indication of a suspicious call. An alternative is to introduce an out-of-band derivative of STIR, but this would inherently be obsolete once end-to-end SIP is available; NICC does not believe that this can be economically justified.

Further, the Interim and Final Stage solutions as set out in this document would require a database of which numbers are assigned for usage on which networks, i.e. a common numbering database. Such databases have considerable cost, and NICC's advice would be that it is unlikely that the cost could be justified solely for STIR-purposes: it is recommended that Ofcom explores whether the functionality could be shared with other applications such as number portability.

Many of the functions associated with the Certificate Authority are outside the traditional knowledge base of NICC, and the challenges of bringing together expertise associated with cryptography/certification and that associated with telephony networks should not be under-estimated. The implementation of SHAKEN could address many of the issues, but there are sufficient differences between the USA and UK markets that even if adoption in the USA is successful, this does not guarantee any form of "off-the-shelf" solution for the UK.

NICC's assessment of the implications for individual networks is that the costs would be significant. It is considered that STIR functionality could be designed into networks as they evolve over the coming decade.

NICC awaits a mandate to develop the associated technical standards for UK networks in readiness for a time where end-end SIP can be expected in the majority of cases.

Annex A (informative): Implementation Phases

This Annex provides a strawman for implementation phases of STIR in the UK

Phase	Description	Pre-requisites	What is signed & by whom	Terminating network actions	What does this give us
Start-up Stage – no numbering database					
1	Start up. Voluntary signing.	Creation of architecture,	NN by originating network. • For PN being Type 1 & 2, Full Attestation. • For PN being Types 3-5. Partial Attestation • For inbound international, Gateway Attestation	N/A	Gets the ball rolling
2a	Terminating network acting upon STIR	Phase 1, end-end SIP		Terminating network uses correctly signed NN to indicate validity of PN, taking into account the level of Attestation	Partial validation of where call was originated
Interim Stage, introduction of database					
2b	Numbering database available	Phase 2a	NN by originating network. • For PN being Type 1 & 2, Full Attestation. • For PN being Types 3-5. Partial Attestation For inbound international, Gateway Attestation	Terminating network uses correctly signed NN – validated against numbering database - to indicate validity of PN, taking into account the level of Attestation	Where signed, confirmation that originator had the rights to use that NN
3	Mandatory signing	Phase 2, regulatory action or industry MoU, all originating networks to be IP			Universal NN signing
4	Mandatory CLI validation	Phase 3, regulatory action or industry MoU, fully IP network		Terminating network uses correctly signed NN to indicate validity of PN and additionally may reject calls based on lack of signing	Full validation of the originating CP, but not necessarily that Type 3-5 PN CLIs are valid.

Phase	Description	Pre-requisites	What is signed & by whom	Terminating network actions	What does this give us
Final Stage Variant I: Originating network PN whitelisting					
5I	Optional originating network validation of PN	Phase 4, process for enterprises to supply full list of PNs to originating network, originating network ability to whitelist CLIs	NN by originating network, but for Types 2, 3 and 5 where PN has been validated, attestation is full rather than partial	As Phase 4	Greater reliability of Types 2, 3 and 5 PNs
6I	Mandatory originating network validation of PN	Phase 5I, regulatory action /industry MoU	NN by originating network, but where Types 2, 3 and 5 PN has been received from customer this must be validated and all CLIs marked as full attestation - partial attestation no longer allowed	Terminating network uses correctly signed NN to indicate validity of PN and additionally may reject calls based on lack of full signing. Only full attestation establishes trust.	Full validation of number to be displayed, (excl Type 4)

Phase	Description	Pre-requisites	What is signed & by whom	Terminating network actions	What does this give us
Variant II: Enterprise signing checked by originating network					
5II	Optional enterprise signing of PN, that is checked by originating network before signing NN	Phase 4, numbering database extended to enterprises	NN by originating network (as above) – if enterprise network has correctly signed PN then on a full attestation basis, otherwise on a partial attestation basis	As Phase 4	Full validation of the originating CP
6II	Enterprise compelled to sign PN and originating network to check this before signing NN. Partial attestation no longer acceptable. Terminating network checks NN but no longer accepts partial attestation	Phase 5II, regulatory action /industry MoU	NN by originating network (as above) – if enterprise network has correctly signed PN then on a full attestation basis.	Terminating network uses correctly signed NN to indicate validity of PN and additionally may reject calls based on lack of full signing. Only full attestation establishes trust.	Full validation of number to be displayed (excl Type 4)

Phase	Description	Pre-requisites	What is signed & by whom	Terminating network actions	What does this give us
Variant III: Enterprise signing checked by terminating network					
5III	Optional enterprise signing of PN flowing through network	Phase 4, numbering database extended to enterprises	Optional enterprise signing of PN, otherwise signing of NN by originating network	As Phase 4	Full validation of the originating CP
6III	Terminating network acts upon signed PNs	Phase 5III, regulatory action /industry MoU	PN signed by enterprise if they're providing it, otherwise signing of NN by originating network on a full attestation basis	Terminating network only trusts calls with either fully attested NN, or signed PN	Full validation of number to be displayed (excl Type 4)

Annex B: (Informative) Changes to call flow to support final stage implementation

The final stages of STIR implementation would change the call flow as set out in Section 6.2, according to which of the variants set out in Section 5 is subsequently adopted.

Variant I – originating network whitelisting
In this approach, the criteria used by the Authentication Service in step four would be that it is able to verify any PN CLI received from the customer against a whitelist of acceptable numbers for that customer. The signing will be marked as “Full attestation”, as the Authentication Service will have verified the PN as legitimate.

Variant II – originating customer signs PN, originating network validates
This approach would require that there is a certificate infrastructure such that originating enterprises/call-centres are able to digitally sign their own PNs. As shown in Figure B.1 below, for calls from such customers, steps 1-6 above would be replaced as follows:

i. The originating SIP terminal creates a SIP INVITE with a telephone number identity in the From: header that, according to customer configuration, may be intended to be used as the PN for display purposes.

ii. The Originating SIP terminal then initiates an originating trigger to the Enterprise Authentication Service for the INVITE.

iii. The Enterprise Authentication Service first determines the legitimacy of the PN CLI being used in the INVITE. If acceptable, the Enterprise Authentication Service then securely requests its private key for the PN from the Enterprise Key Store. If unacceptable then the Authentication Service either fails the call, or passes the call with the Presentation Number unsigned.

iv. The originating SIP terminal, which is registered and authenticated to the Originating Call Server, creates a SIP INVITE with the signed PN.

v. The Originating Call Server populates the P-Asserted-Identity header field asserting the Network Number CLI of the originating SIP Terminal’s network termination point, the Presentation Number CLI received from the SIP terminal in the From: header field and the signed PN. The Originating Call Server then initiates an originating trigger to the Authentication Service for the INVITE.

vi. The Authentication Service first determines the legitimacy of the PN CLI being used in the INVITE; this would be by carrying out the functions of a Verification Service for the signed PN.

    a. It uses the “info” parameter information in the Identity header field as specified in RFC8224 to determine the Certificate Store Uniform Resource Identifier (URI) and makes an HTTPS request to the Certificate Store.
    b. It validates the certificate then extracts the public key. It constructs the RFC8224 format and uses the public key to verify the signature in the Identity header field, which validates that the Presentation Number CLI used is authentic. Note that in order to do this, a common numbering database at the level of providing the enterprise to which numbers are assigned would be required.

vii. If acceptable, the Authentication Service then securely requests its private key for the Network Number from the Key Store. If unacceptable then the Authentication Service either fails the call, passes the call with no signed Network Number, or inserts a suitable Presentation Number (which may be a copy of the Network Number) and then securely requests its private key for the Network Number from the Key Store.

viii. The Key Store provides the private key in the response, and the Authentication Service signs the INVITE and adds an Identity header field as specified in RFC8224 using the Network Number CLI in the P-Asserted-Identity header field. The signing will be marked as “Full attestation”.

ix. The Authentication Service passes the INVITE to the Originating Call Server.

x. The Originating Call Server, through standard resolution, routes the call to the egress Border Gateway (NB this would contain the signed Network Number, but the signed Presentation Number would be discarded).

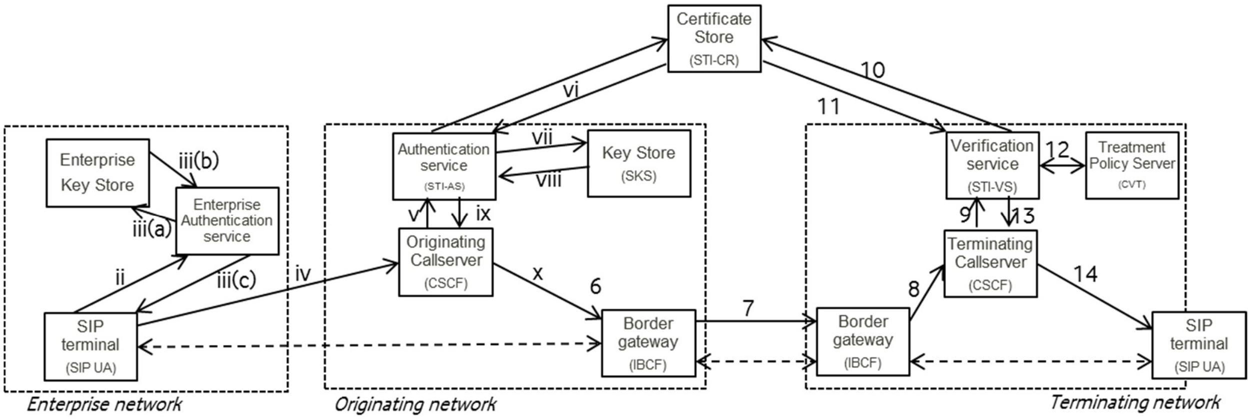

Figure B.1: Final Stage Variant II

<u>Variant III – originating customer signs PN, Terminating Network validates</u>
This approach would require that there is a certificate infrastructure such that originating enterprises/call-centres are able to digitally sign their own PNs. It should be noted that it is an open issue whether, for this variant the Originating Network would sign the NN in addition to the customer signing the PN; for the purpose of this description, it is assumed that they would not. As shown in Figure B.2, for calls from such customers, steps 1-6 above would be replaced as follows:

i. The originating SIP terminal creates a SIP INVITE with a telephone number identity in the From: header that, according to customer configuration, may be intended to be used as the PN for display purposes.

ii. The Originating SIP terminal then initiates an originating trigger to the Enterprise Authentication Service for the INVITE.

iii. The Enterprise Authentication Service first determines the legitimacy of the PN CLI being used in the INVITE. If acceptable, the Enterprise Authentication Service then securely requests its private key for the PN from the Enterprise Key Store. If unacceptable then the Authentication Service either fails the call, or passes the call with the Presentation Number unsigned.

iv. The originating SIP terminal, which is registered and authenticated to the Originating Call Server, creates a SIP INVITE with the signed PN.

v. The Originating Call Server populates the P-Asserted-Identity header field asserting the Network Number CLI of the originating SIP Terminal's network termination point, the Presentation Number CLI received from the SIP terminal in the From: header field and the signed PN.

vi. The Originating Call Server, through standard resolution, routes the call to the egress Border Gateway.

In Steps 10 and 11, the certificate concerned would relate to the enterprise customer rather than the Originating Network, and the Verification Service would act upon the signed PN rather than NN. Note that in order to do this, a common numbering database at the level of providing the enterprise to which numbers are assigned would be required.

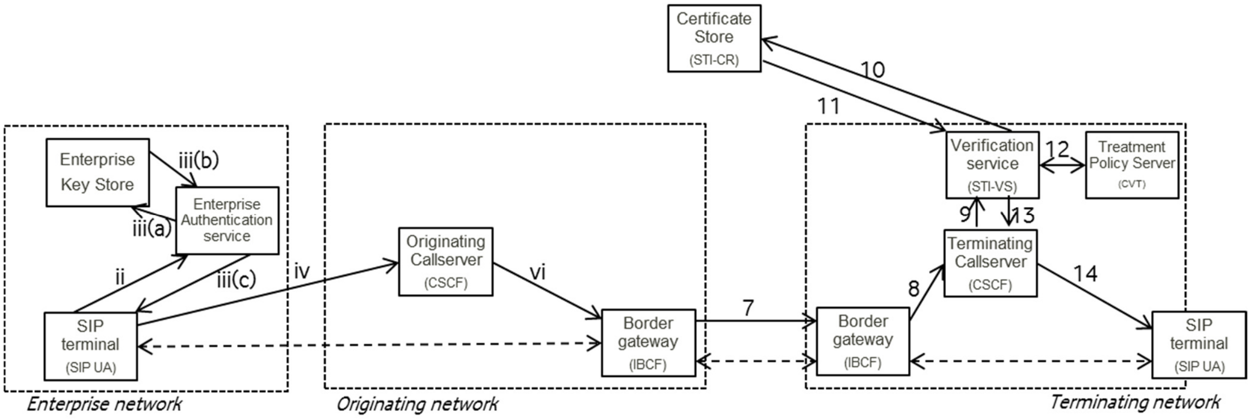

Figure B.2: Final Stage Variant III

Annex C (informative): Approaches considered for credential distribution

As set out in Section 7, a series of approaches were considered for the distribution of credentials in a UK STIR implementation, with it being concluded that Approach 2 best meets the identified needs;

- That the approach does not preclude early implementation
- That the approach allows identification of the originating network;
- That the approach confirms that the originating network had the rights to use the CLI;
- That the approach as far as possible is similar to that used in other jurisdictions, in order to minimise the chances of requiring UK-specific equipment;
- That the signalling network will not be compromised;
- That the solution is secure against man-in-the-middle attacks;
- That it is preferable not to have to create new central (i.e. third party) functions;
- That post-dial-delay is minimised;
- That costs are minimised;
- That caching of information is facilitated to minimise external network queries;
- That it could be scalable to cover final stage variants (see Section 4)

This Annex describes the alternative approaches considered, and why Approach 2 was considered superior to the others.

Approach 1

The first approach considered is depicted in Figure C.1. This approach, which is similar to that adopted in the launch phase by the US SHAKEN initiative, establishes only where the call was originated, not that the originating network necessarily has the rights to use that CLI.

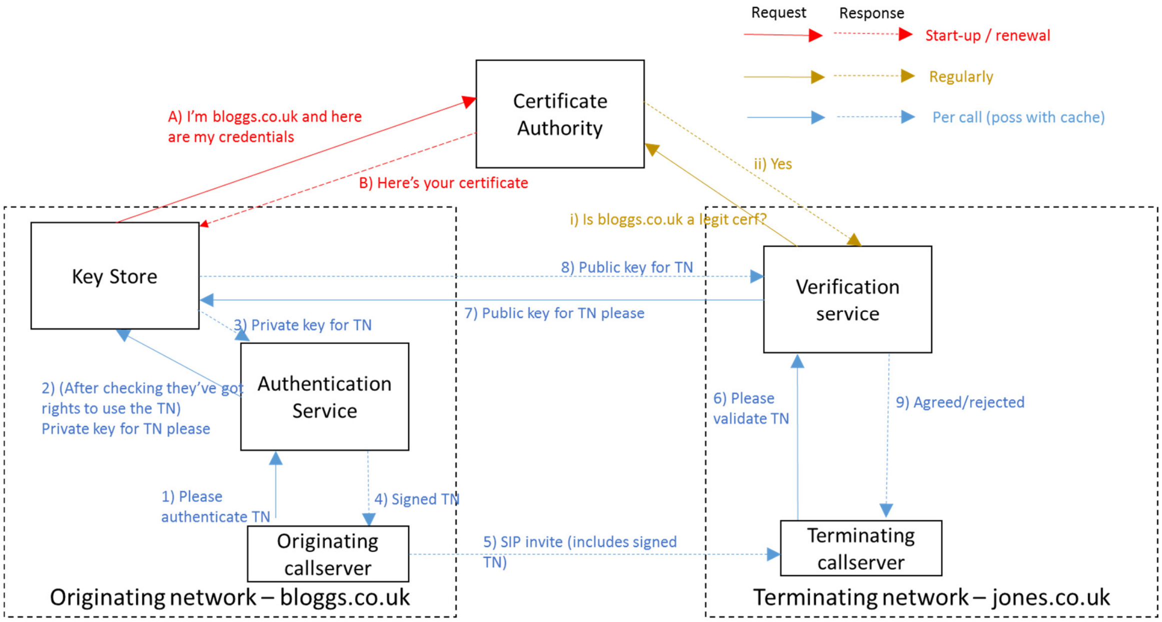
Request
Response
Start-up / renewal
Regularly
Per call (poss with cache)
Certificate Authority
A) I'm bloggs.co.uk and here are my credentials
B) Here's your certificate
ii) Yes
i) Is bloggs.co.uk a legit cert?
Key Store
8) Public key for TN
7) Public key for TN please
3) Private key for TN
Verification service
Authentication Service
2) (After checking they've got rights to use the TN) Private key for TN please
6) Please validate TN
9) Agreed/rejected
1) Please authenticate TN
4) Signed TN
Originating callserver
5) SIP invite (includes signed TN)
Terminating callserver
Originating network – bloggs.co.uk
Terminating network – jones.co.uk

Figure C.1: Approach 1

As part of a start-up/renewal phase, the originating network would present its credentials to a central Certificate Authority, and get a certificate from which it can generate private and public keys for its numbers. On each call, the Authentication Service would then sign the CLI using the private key, and this information would be passed in the SIP signalling. At the terminating network, the Verification Service would check that the signing certificate is legitimate (as this information would be on a per-originating network basis, this legitimacy information could be readily cached), and retrieve the public keys to decrypt the STIR information from the originating network.

Approach 1 is considerably simpler than the subsequent approaches considered, as it does not require a central database (CDB) of which networks have the rights to use which numbers; conversely it only serves to identify which network originated the call, not whether they had the right to do so with the CLI in question – this is a key requirement for any UK implementation of STIR. This limitation could be considered reasonable as a start-up phase to achieve rapid implementation, given the complexity in establishing a CDB.

Approach 1a

Approach 1a is a variant of Approach 1 and is depicted in Figure C.2. This approach differs from Approach 1 only in that the public key information is conveyed in the SIP signalling rather than having to be retrieved by the Verification Service: this has advantages in reducing the scope for post-dial delay by removing steps 7) and 8) in Approach 1.

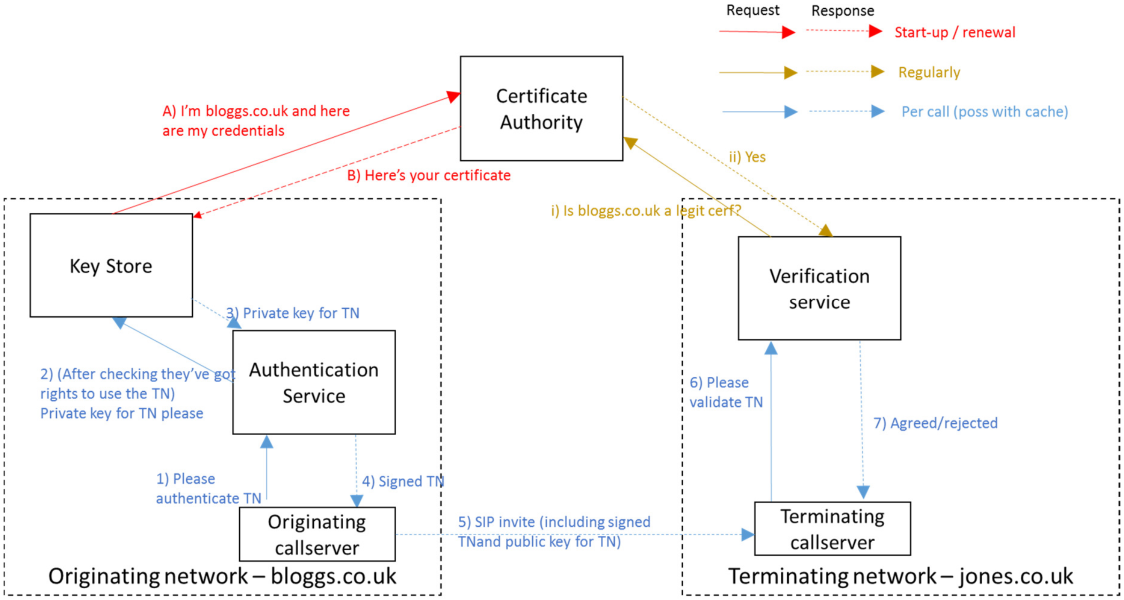
Request
Response
Start-up / renewal
Regularly
Per call (poss with cache)
Certificate Authority
A) I'm bloggs.co.uk and here are my credentials
B) Here's your certificate
ii) Yes
i) Is bloggs.co.uk a legit cert?
Key Store
3) Private key for TN
Authentication Service
2) (After checking they've got rights to use the TN) Private key for TN please
1) Please authenticate TN
4) Signed TN
Originating callserver
5) SIP invite (including signed TNand public key for TN)
Originating network – bloggs.co.uk
Verification service
6) Please validate TN
7) Agreed/rejected
Terminating callserver
Terminating network – jones.co.uk

Figure C.2: Approach 1a

There are, however, concerns about carrying the public key within the signalling. Firstly, this will significantly increase the size of the signalling headers that need to be processed by call servers in the call path, however it has not been possible to quantify this. Secondly, the one reference implementation that exists for STIR – SHAKEN in the USA – does not adopt the in-band carriage of the public key. It is of course possible that subsequent implementations will adopt that approach, but NICC considered it prudent to follow the same basic approach as the sole implementation of STIR.

Therefore, Approach 1a was rejected as a preferred approach as it did not fulfil the goal of verifying that the originating network had the right to use the CLI, and risked requiring a UK-specific implementation.

Approach 2

This option, depicted in Figure C.3, builds upon Approach 1 to build in a check that the originating network actually has the right to use the number concerned as a CLI. To do this a CDB is required, and this is queried by the terminating network Verification Service either in series with requesting the public key to check the signed CLI, or in parallel with it.

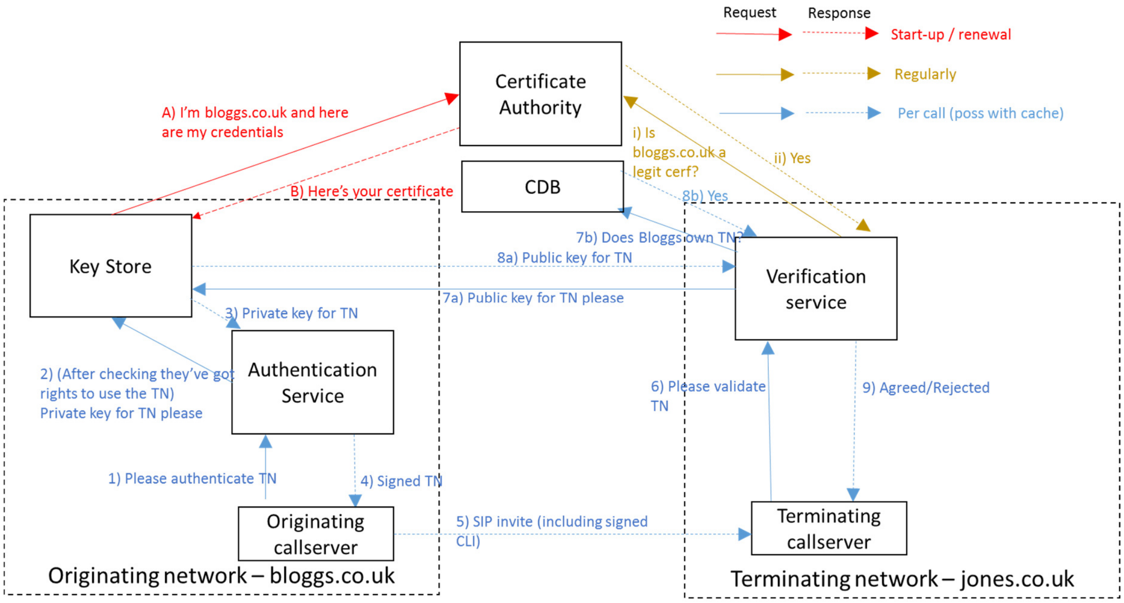
Request
Response
Start-up / renewal
Regularly
Per call (poss with cache)
Certificate Authority
CDB
A) I'm bloggs.co.uk and here are my credentials
B) Here's your certificate
i) Is bloggs.co.uk a legit cerf?
ii) Yes
8b) Yes
7b) Does Bloggs own TN?
8a) Public key for TN
7a) Public key for TN please
Key Store
3) Private key for TN
2) (After checking they've got rights to use the TN) Private key for TN please
Authentication Service
1) Please authenticate TN
4) Signed TN
Originating callserver
5) SIP invite (including signed CLI)
Verification service
6) Please validate TN
9) Agreed/Rejected
Terminating callserver
Originating network – bloggs.co.uk
Terminating network – jones.co.uk

Figure C.3: Approach 2

This approach does fulfil the requirement to confirm that the originating network is permitted to use the CLI, and has the advantage that the operation of the CDB can be detached from the Certificate Authority, thus allowing the optimal supplier for each to be chosen. Approach 2 also facilitates an evolutionary approach from Approach 1, with the CDB check being introduced once the latter was available. However, set against this it either requires two checks at the Verification Service – with associated potential post-dial delay. Although Approach 2 wasn't favoured when NICC initially examined the approaches, on balance it is now considered the best approach.

Approach 2a

Approach 2a is a variant of Approach 2 and is depicted in Figure C.4. This approach differs from Approach 2 only in that the public key information is conveyed in the SIP signalling rather than having to be retrieved by the Verification Service: this has advantages in reducing the scope for post-dial delay by removing steps 7a) and 8a) in Approach 2. This removes the issue with Approach 2 of needing either sequential queries (hence possible post-dial delay) or parallel queries (hence complexity) to be made.

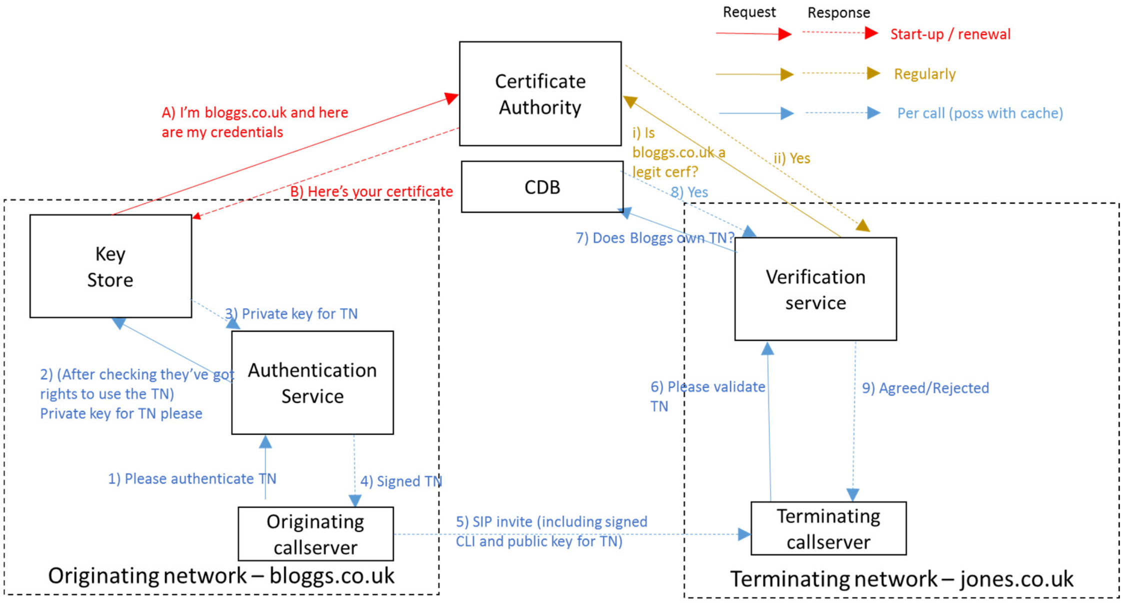
Request
Response
Start-up / renewal
Regularly
Per call (poss with cache)
Certificate Authority
A) I'm bloggs.co.uk and here are my credentials
B) Here's your certificate
CDB
i) Is bloggs.co.uk a legit cerf?
ii) Yes
8) Yes
7) Does Bloggs own TN?
Key Store
3) Private key for TN
Authentication Service
2) (After checking they've got rights to use the TN) Private key for TN please
1) Please authenticate TN
4) Signed TN
Originating callserver
5) SIP invite (including signed CLI and public key for TN)
Verification service
6) Please validate TN
9) Agreed/Rejected
Terminating callserver
Originating network – bloggs.co.uk
Terminating network – jones.co.uk

Figure C.4: Approach 2a

As with Approach 1a, however, NICC was concerned at following a fundamentally different approach to that adopted in SHAKEN. Therefore, this approach was not adopted as NICC's preferred option.

Approach 3

Approach 3 differs from the other approaches in that it centralises the distribution of public keys: as these are held centrally, they would only be accepted and distributed to terminating network Verification Services if they were deemed (by reference to a CDB) to be valid for the originating network concerned. The approach is illustrated in Figure C.5.

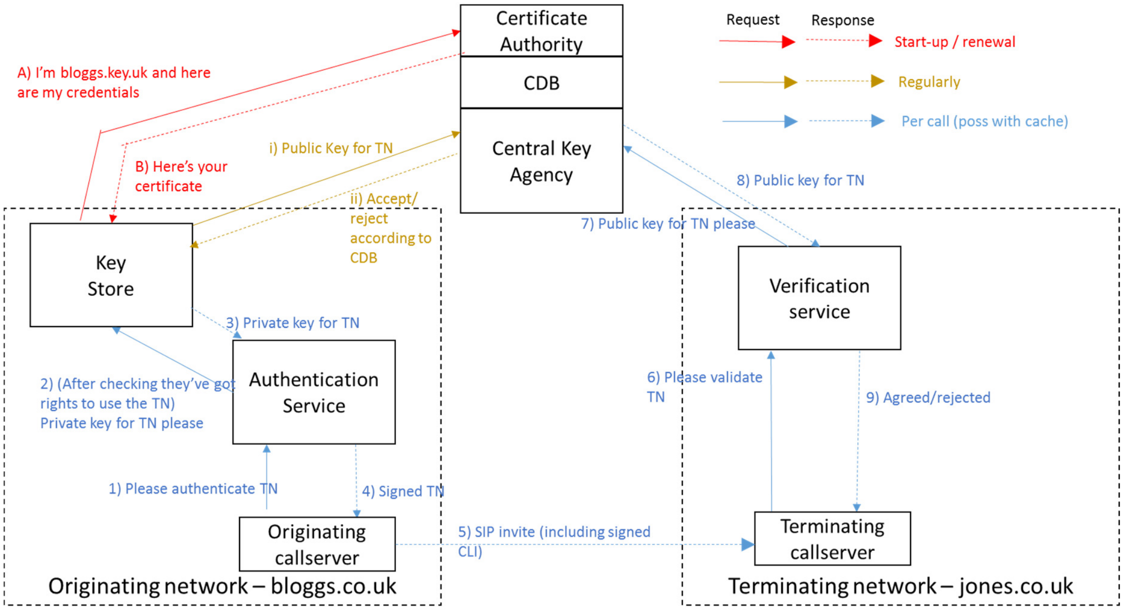
Certificate Authority
CDB
Central Key Agency
Request
Response
Start-up / renewal
Regularly
Per call (poss with cache)
A) I'm bloggs.key.uk and here are my credentials
B) Here's your certificate
i) Public Key for TN
ii) Accept/ reject according to CDB
8) Public key for TN
7) Public key for TN please
Key Store
3) Private key for TN
2) (After checking they've got rights to use the TN) Private key for TN please
Authentication Service
1) Please authenticate TN
4) Signed TN
Originating callserver
Originating network – bloggs.co.uk
5) SIP invite (including signed CLI)
Verification service
6) Please validate TN
9) Agreed/rejected
Terminating callserver
Terminating network – jones.co.uk

Figure C.5: Approach 3

Like Approach 2, Approach 3 has the advantage that it not only allows identification of the originating network, but also verification of whether it had the right to use the CLI in question. Additionally, as the public keys are held in a single location, this would assist in bulk downloading all of them in order that on a per-call basis the Verification Services could act autonomously within their own network operator domain. Furthermore, only a single query is needed, thus easing implementation at the terminating Verification Service.

Set against this, however, it places more functionality into a central body, which could increase costs of a monopoly/shared function. It was therefore concluded that whilst this was an approach that would work, it was not preferred.

Approach 4

Approach 4 attempts to slim down the functions carried out by the central agency by restricting it to storing pointers to the location of the public keys, rather than the keys themselves. The information conveyed in the SIP signalling would thus always point to a Central Key Agency rather than directly to the originating network, and the Central Key Agency would only refer queries on to the originating network if it was valid for that CLI. Figure C.6 illustrates the approach.

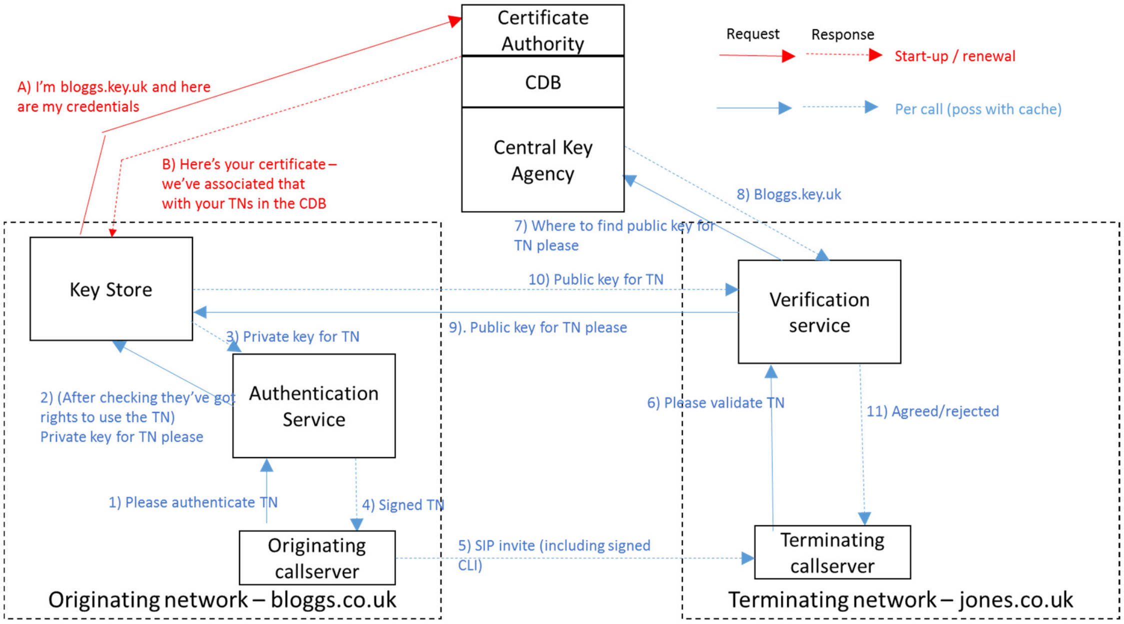
Certificate Authority
CDB
Central Key Agency
Request
Response
Start-up / renewal
Per call (poss with cache)
A) I'm bloggs.key.uk and here are my credentials
B) Here's your certificate – we've associated that with your TNs in the CDB
8) Bloggs.key.uk
7) Where to find public key for TN please
Key Store
10) Public key for TN
9). Public key for TN please
Verification service
3) Private key for TN
Authentication Service
2) (After checking they've got rights to use the TN) Private key for TN please
6) Please validate TN
11) Agreed/rejected
1) Please authenticate TN
4) Signed TN
Originating callserver
5) SIP invite (including signed CLI)
Terminating callserver
Originating network – bloggs.co.uk
Terminating network – jones.co.uk

Figure C.6: Approach 4

Whilst this approach does address the issue of minimising the central functions when compared to Approach 3, conversely like Approach 2 it requires two queries – and in this case they must be sequential, thus raising concerns about the impact upon post dial delay. For this reason, Approach 4 was not selected as NICC's preferred option.

Approach 5

Under Approach 5, there would be a central Certificate Authority for the UK numbering plan, which would distribute the certificates that are used to generate public and private keys for usage in STIR. Certificates would only be distributed for the numbers that a given originating network is permitted to use, and likewise the Certificate Authority would publish the numbers that are valid for a given certificate to terminating network Verification Services. The approach is illustrated in Figure C.7 below.

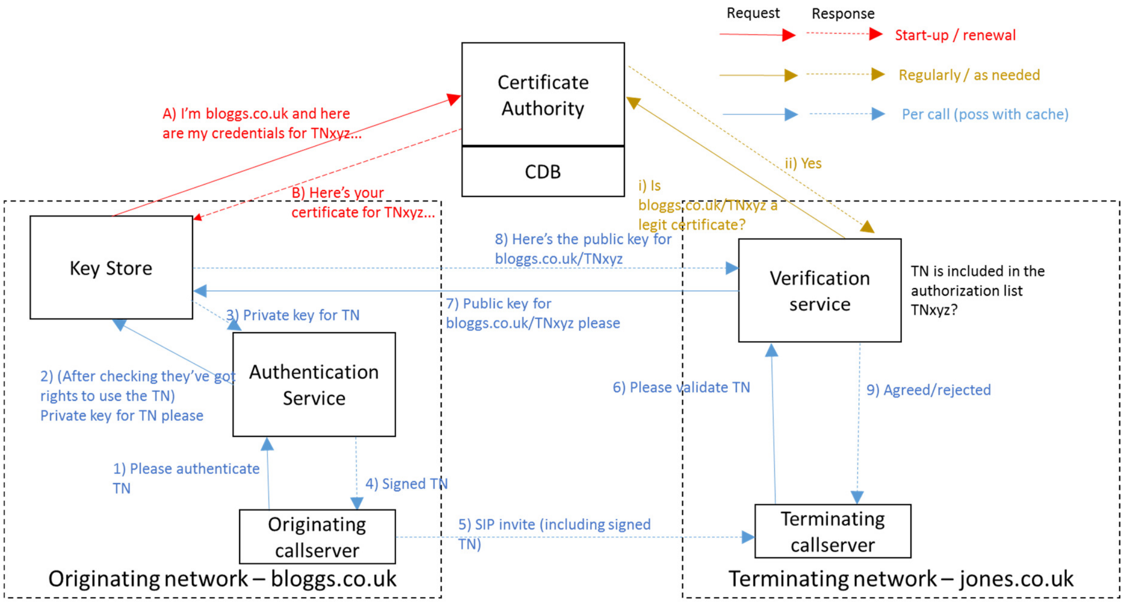
Request
Response
Start-up / renewal
Regularly / as needed
Per call (poss with cache)
Certificate Authority
CDB
A) I'm bloggs.co.uk and here are my credentials for TNxyz...
B) Here's your certificate for TNxyz...
i) Is bloggs.co.uk/TNxyz a legit certificate?
ii) Yes
Key Store
8) Here's the public key for bloggs.co.uk/TNxyz
Verification service
TN is included in the authorization list TNxyz?
7) Public key for bloggs.co.uk/TNxyz please
3) Private key for TN
Authentication Service
2) (After checking they've got rights to use the TN) Private key for TN please
6) Please validate TN
9) Agreed/rejected
1) Please authenticate TN
4) Signed TN
Originating callserver
5) SIP invite (including signed TN)
Terminating callserver
Originating network – bloggs.co.uk
Terminating network – jones.co.uk

Figure C.7: Approach 5

With proper management of the mapping of telephone numbers to certificates/keys, when NICC first examined STIR, it was considered that Approach 5 offered the best way of achieving the twin goals of identifying both the originating network and that they had the right to use the CLI in question. So long as there were sufficient volumes of numbers associated with each certificate, the volume of requests about certificates to the Certificate Authority could be managed to the point of it being possible to bulk download or cache the data so that per-call queries aren't required. Similarly, although in principle stages 7) and 8) in Figure C.7 imply a query from the terminating to originating network, with suitable management of number-to-key mapping this information could largely be cached, hence reducing the volume of these queries.

However, Approach 5 involves setting up a Certificate Authority with intimate knowledge of the UK numbering plan, including a common numbering database, before any calls can be signed. On review, NICC has concluded that this would stifle the introduction of STIR, so it is no longer favoured.

Approach 5a

The final option considered was Approach 5a, which is essentially Approach 5 but with the public key information being carried within the SIP signalling, as depicted in Figure C.8.

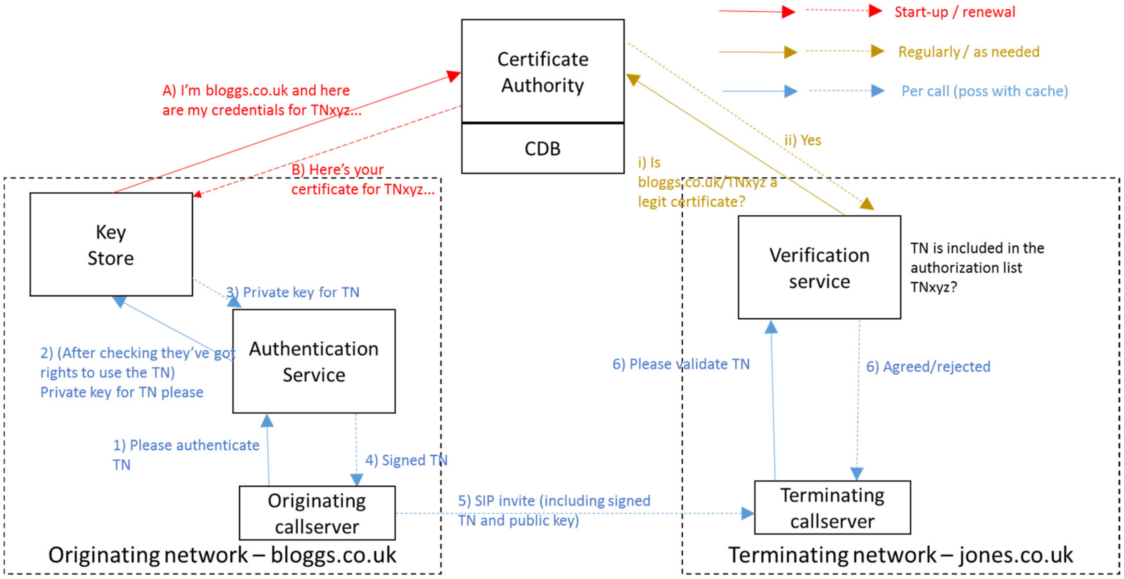
Start-up / renewal
Regularly / as needed
Per call (poss with cache)
Certificate Authority
CDB
A) I'm bloggs.co.uk and here are my credentials for TNxyz…
B) Here's your certificate for TNxyz…
i) Is bloggs.co.uk/TNxyz a legit certificate?
ii) Yes
Key Store
3) Private key for TN
2) (After checking they've got rights to use the TN) Private key for TN please
Authentication Service
1) Please authenticate TN
4) Signed TN
Originating callserver
Originating network – bloggs.co.uk
5) SIP invite (including signed TN and public key)
Verification service
TN is included in the authorization list TNxyz?
6) Please validate TN
6) Agreed/rejected
Terminating callserver
Terminating network – jones.co.uk

Figure C.8: Approach 5a

This option in principle removes stages 7) and 8) of Approach 5, but as has been discussed, it is likely that these will be replaced by reference to cached data in any case. Further, by carrying the public key information in signalling, this increases the size of the SIP signalling and potentially puts the UK out of step with international implementations. NICC therefore rejected this option.

Annex D (informative): Bibliography

In addition to the documents set out in the Reference section, the following documents provide useful background reading to STIR an SHAKEN

- IETF RFC7340: "Secure Telephone Identity Problem Statement and Requirements"
- IETF RFC8225: "PASSportT: Personal Assertion Token"
- IETF RFC8226: "Secure Telephone Identity Credentials: Certificates"
- draft-ietf-stir-oob-07 (3rd March 2020): "STIR Out of Band Architecture and Use Cases"
- draft-ietf-stir-cert-delegation-02 (3rd March 2020): "STIR Certificat Delegation"
- ATIS-1000080.v002: "Joint ATIS/SIP Forum Standard – Signature-based Handling of Asserted information using toKENs (SHAKEN): Governance Model and Certificate Management".
- ATIS-1000082: "Technical Report on SHAKEN APIs for a Centralized Signing and Signature Validation Server"
- ATIS-1000084: "Technical Report on Operational and Management Considerations for SHAKEN STI Certification Authorities and Policy Administrators"
- ATIS-0300116 (October 2019): "Interoperability Standards between Next Generation Networks (NGN) for Signature-Based Handling of Asserted information Using Tokens (SHAKEN)"
- ITU-T X.509 (October 2016): "Information technology - Open Systems Interconnection - The Directory: Public-key and attribute certificate frameworks"
- ETSI TS 103 486: "Identity management and naming schema protection mechanisms" (not yet published)

History

Document history		
Version	Date	Milestone
1.1.1	18 April 2018	Initial publication
2.1.1	10th September 2020	Second published version