# NICC ND 1522 V1.1.1 (2018-04)

*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

NOTICE OF COPYRIGHT AND LIABILITY



The present document may be made available in more than one electronic version or in print. In any case of existing or perceived difference in contents between such versions, the reference version is the Portable Document Format (PDF). In case of dispute, the reference shall be that printing on NICC printers of the PDF version kept on a specific network drive within the NICC.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other NICC documents is available at:

http://www.niccstandards.org.uk/publications/

If you find errors in the present document, please send your comments to:

mailto:help@niccstandards.org.uk



If you have any comments concerning the accuracy of the contents of this document, please write to:

The Technical Secretary,
NICC Standards Ltd,
PO Box 3090
Eastbourne,
BN21 9HA

# Contents

# 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.

# Introduction

UK consumers receive large volumes of unsolicited and unwelcome marketing calls. Although some calls display a valid 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.

The Internet Engineering Task Force (IETF) has standardised a mechanism to digitally sign identities (such as CLIs) in order that terminating ends 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 (05/97) "The international public telecommunication numbering plan"

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

[6] ND1016 V3.2.1 (2015-03) Requirements on Communications Providers in relation to Customer Line Identification display services and other related services

[4] ND1035 V2.1.1 (2016-02): SIP Network to Network Interface Signalling

[5] draft-ietf-stir-rfc4474bis-16 (9th February 2017): "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] ND1034 V1.1.1 (2017-03) 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 4 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 4 of this document, the status where the entity signing the CLI can unequivocally assert that the originator of the call has rights to the CLI.

**Gateway Attestation:** As set out in Section 4 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 initial implementation of UK STIR whereby the Network Number would be signed, potentially on a Partial Attestation basis, as set out in Section 4 of this document.

**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 4 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 not that the customer necessarily has the rights to use that number.

**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].

**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** |
|---|---|---|
| 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) |
| Business line where the customer has not requested a Presentation Number | | |
| 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) |
| 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 |
| 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 |
|---|---|---|
| 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. |
| 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. |
| 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 |
| 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. |
| UK mobile calling from their home network (i.e. not roaming)<br>UK mobile roaming overseas (with home network routing enabled)<br>UK mobile roaming overseas (with direct routing enabled) | 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. |
| 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) | |
| Overseas call-centre wishing to display UK number: connected to UK network (i.e. 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 |
|---|---|---|
| Overseas call-centre 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 draft RFC4474bis [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 infrastructure to distribute the private keys associated with digital signing (see Section 7): the delegation path would need to be via the end-customer to their chosen originating network operators. 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. Following this approach, signing the P-A-ID header field 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. The remainder of this document terms this approach as the “Interim Stage”.

As set out in Section 6.2, a parameter within STIR allows 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". 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

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 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.

Even within the proposed 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 Architecture

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

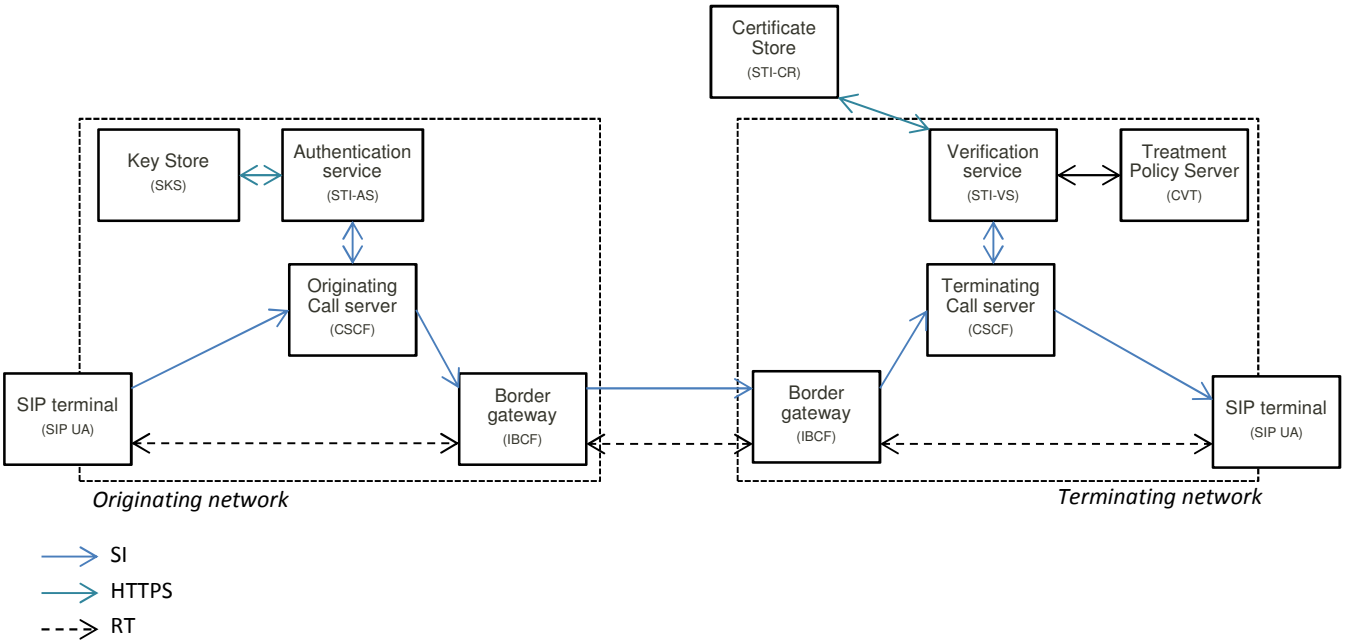


**Figure 6.1.a:** STIR architecture

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 RFC4474bis. 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 RFC4474bis. 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 interim stage (as described in Section 5), where CLIs are authenticated by the originating network via STIR.

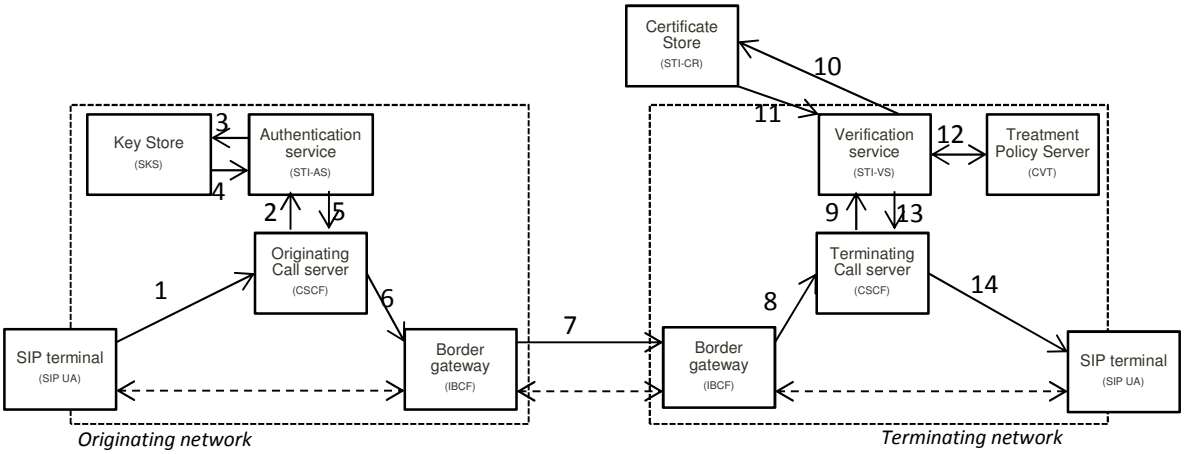


**Figure 6.2.a:** STIR call flow

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.

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

3. The Authentication Service first determines the legitimacy of the PN CLI being used in the INVITE (during the Interim Stage, 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 no signed 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.

4. 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 draft-ietf-stir-rfc4474bis 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".
5. The Authentication Service passes the INVITE to the Originating Call Server.
6. The Originating Call Server, through standard resolution, routes the call to the egress Border Gateway.
7. The INVITE is routed over the NNI through the standard inter-domain routing configuration.
8. The Terminating Network's ingress Border Gateway receives the INVITE over the NNI.
9. The Terminating Call Server initiates a terminating trigger to the Verification Service for the INVITE.
10. The Terminating Verification Service uses the "info" parameter information in the Identity header field as specified in draft-ietf-stir-rfc4474bis to determine the Certificate Store Uniform Resource Identifier (URI) and makes an HTTPS request to the Certificate Store.
11. The Verification Service validates the certificate then extracts the public key. It constructs the draft-ietf-stir-rfc4474bis format and uses the public key to verify the signature in the Identity header field, which validates that the Network Number CLI used is authentic. 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.
12. 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.
13. 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.
14. 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<br><br>Route call, with an advice whisper on answer<br><br>Route call with a visual advice of un-validated PN (for IP phones with such capability)<br><br>Send call to voicemail box<br><br>Block call (based on user opt-in to blocking)<br><br>Send call through a screening service before routing<br><br>Block call (based on CP choice dependent on upstream carrier) |
| Partial | Route call with or without displayed PN<br><br>Route call, with a cautionary advice whisper on answer<br><br>Route call with a visual advice of low PN reliability score or 'trust marking' (for IP phones with such capability)<br><br>Send call to voicemail box<br><br>Block call (based on user opt-in to blocking)<br><br>Send call through a screening service before routing<br><br>Block call (based on CP trust level of originating CP) |
| Full | Route call with displayed PN<br><br>Route call with visual advice of high PN reliability score or 'trust marking' (for IP phones with such capability)<br><br>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 | Conditions resulting in full trust |
|---|---|---|---|
| Interim Stage (phase 2) | N/A | a. No signing of NN CLI, or<br>b. NN CLI is signed with gateway attestation, or<br>c. NN CLI is signed with partial attestation and originator is untrusted | a. NN CLI is signed with full attestation<br>b. NN CLI is signed with partial attestation and originator is trusted |
| Interim Stage (phase 4) – Note 2 | No signing of NN CLI | a. NN CLI is signed with gateway attestation, or<br>b. NN CLI is signed with partial attestation and originator is untrusted | a. NN CLI is signed with full attestation<br>b. NN CLI is signed with partial attestation and originator is trusted |
| Final Stage (phase 6I/6II) – Note 3 | a. No signing of NN CLI, or<br>b. NN CLI is signed with partial attestation (Note 4) | NN CLI is signed with gateway attestation | NN CLI is signed with full attestation |
| Final Stage (phase 6III) – Note 5 | a. No signing of either PN or NN CLI<br>b. NN CLI signed with partial attestation (Note 4) | NN CLI is signed with gateway attestation | a. NN CLI is signed with full attestation, or<br>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. By this time the expectation is that all Network Number CLIs would be signed.
3. 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 II).
4. The terminating network could optionally treat this case as having partial trust according to the level of trust they have in the originating CP.
5. Under Final Stage Variant III, 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. The most promising option for UK implementation is Approach 5, which is set out in Figure 7.a below.

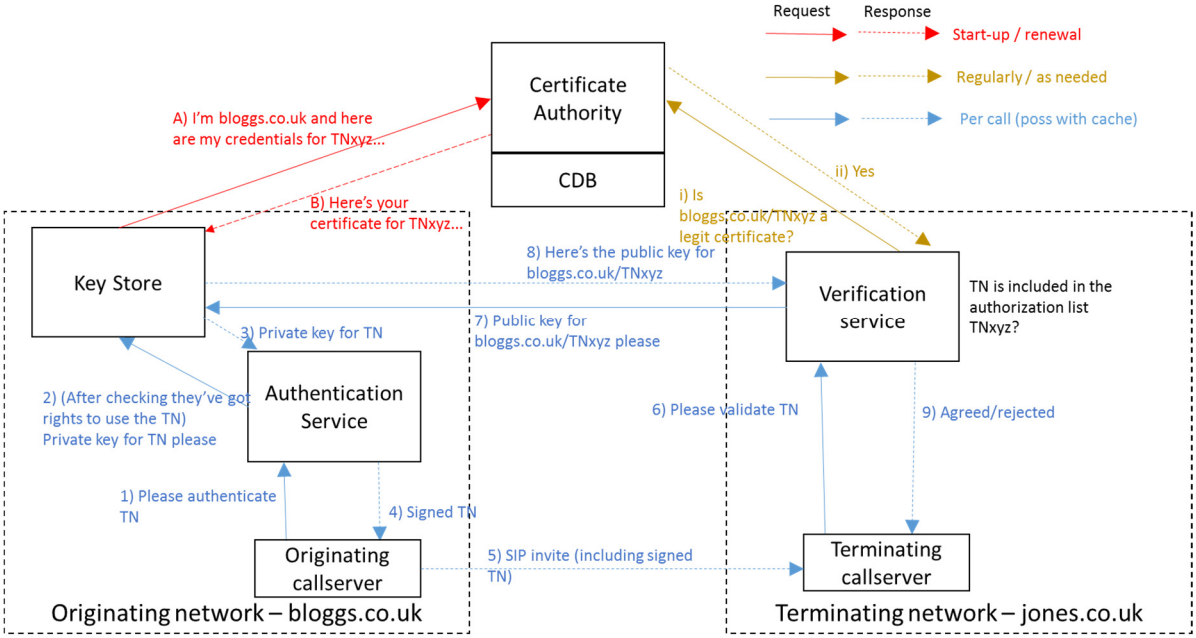


**Figure 7.a:** Recommended STIR credentials treatment for UK

Under this 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 have a record of which numbers are assigned for usage by which originating networks, to ensure that certificates are only distributed to those permitted to use them. As the CLIs used by originating networks can be subject to number portability, inherently this means that the record of number assignments will need to be at an individual number level, i.e. a Central Database of individual numbers (CDB). It is recognised that this would add considerable cost and complexity to STIR implementation, but without that database, there is no way of assessing that an originating network has the rights to sign a given CLI (to phrase this a different way, STIR without a CDB will identify which network originates a call, but not whether the CLI used on that call is one for which it has rights).

The recommended approach would work as follows:

Start-up/renewal phase
When an originating network wishes to participate in STIR, it would communicate with the Certificate Authority, providing a list of telephone numbers which it wishes to associate with a given certificate. The Certificate Authority would validate that the numbers are all within ranges assigned to the originating network, or are numbers ported to that originating network. If so, it would issue a certificate to be used to generate public and private keys for that list of telephone numbers.

A slightly different implementation option that could be used is that the Certificate Authority simply makes available a series of certificates to the originating network, based upon the number ranges that the originating network has been assigned and numbers that it has imported.

The originating network would then generate public and private keys to be used for the list of numbers: the private keys would be kept securely (the Key Store function in Figure 6.2.a), and the public keys made available to terminating networks (the Certificate Store function in Figure 6.2.a).

The originating network would need to periodically renew the certificates associated with a given set of numbers. The frequency of this renewal is for further study, but would be driven by two factors, namely ensuring that it is not possible to reverse engineer the private keys for a given certificate, and ensuring that the certificate has not gone "stale" in terms of customer churn meaning it contains many numbers which are no longer valid for that originating network. Renewal of the certificate would essentially be a re-application for a new certificate.

Pre-provisioning of the terminating network Verification Service functions
The Certificate Authority would also make available a list of all of the certificates issued to originating networks. Terminating networks could therefore proactively download a list of certificates and the telephone numbers to which they apply.

Per-call operation
The per-call operation is as set out in steps 1) to 8) in Figure 7.a. The Authentication Service in the originating network would sign the CLI using the private key, and this information would be conveyed to the terminating network. The Verification Service would then check that the certificate is valid for the telephone number concerned, and retrieve the public keys. Although this retrieval would in principle be from the originating network, in practical terms the information would only need to be retrieved once for the certificate concerned, and could be cached thereafter.

Grouping of numbers to certificates
Grouping of numbers within certificates would represent one of the greatest technical challenges associated with STIR, because the nature of certificates being valid for a series of numbers would be a departure from the more industrialised current arrangements of certificates being valid for internet domain names. Certificates would be valid for a series of numbers – in some cases individual (for example ported numbers), in other cases contiguous ranges of numbers. There is a balance to be struck in deciding how many telephone numbers would be grouped within a given certificate to be signed by a given key set: too many numbers and an individual certificate will constantly need amending as numbers churn between network operators; too few and there would be a large volume of certificates to manage, with little prospect of caching key information.

The optimal approach may vary between network operators. For example, a large-scale national operator might have a certificate representing each geographic area code. Conversely, a small

network operator with limited volumes of numbers could have a single certificate covering all of their operations. Although superficially an individual network operator matter, in fact there is (in economic terms) an externality in that the decisions taken by the originating network operator of how they group their numbers into certificates will impact upon the terminating networks' Verification Service functions. Therefore, it is recommended that rules will be agreed for this grouping, which could vary according to the size of the originating network. In the model described above whereby the Certificate Authority sets the mapping of numbers to certificates, these rules would be implicitly imposed.

Churn of numbers – adding new numbers, porting numbers

Thought will need to be given to how numbers are added into and removed from certificates once STIR is in-life. Although this might be considered an operational issue, in reality it has profound implications for how often certificates need re-issuing, hence the costs involved in establishing a Certificate Authority.

An originating network could choose to associate all of their numbers with certificates, or alternatively just those which are in active service. If the latter approach is taken, this implies that associating the number with a certificate becomes part of the provisioning process. This could result in a large volume of changes to certificates, because for example in the "one certificate per area code" model outlined above, the certificates for every area code which had new numbers assigned to end-users would need to be refreshed to include those new numbers. A better approach could be for each operator to have a "churn" certificate, into which new numbers are added. Periodically – for example weekly or monthly – those numbers could then be swept into the correct certificate for the number concerned.

Similar considerations apply with number portability. In principle, when a number is ported it should be removed from the "donor" certificate and added to the "recipient" certificate in lockstep with the porting exercise. However, in practical terms this can be avoided. If the number is left in the donor certificate until such a time that this certificate is routinely refreshed (weekly, monthly), there is limited risk introduced; the risk is merely that the donor operator could make calls using that CLI in the meantime. On the recipient side, there is a need to get the ported number into a certificate as soon as possible, because it would otherwise not be possible to sign it for use as a CLI. Like the new number case above, though, it will be possible to put the number into a "churn" certificate – possibly in advance of the number port itself - and sweep it into its ultimate certificate on a periodic basis.

Inbound international calls

Whilst the above approach 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 (which may not be restricted to the 07 range), 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 that is already signed*
In this situation it is recommended that the gateway node would pass the signed CLI through transparently, and the terminating network then seeks 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. It should be noted that the P-A-ID header field may contain a number that (pre-STIR) would have been discarded by the international gateway as unreliable; under this model that P-A-ID header field would now be 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 which would allow the international gateway to be traced, but is considered preferable to substituting the P-A-ID header field and hence throwing away the signed information.

*Inbound call contains an unsigned P-A-ID header field but the gateway has no justification to distrust it*
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 gateway certification (see Section 6.3). As the gateway network could in principle be receiving calls with *any* P-A-ID header field, it would not be possible to use a certificate (hence keys) associated with particular numbers, so the one issued by the Certificate Authority would be valid for any number, and the check from the Verification Service to the Certificate Authority depicted in Figure 7.a would merely verify that the gateway network is known by the Certificate Authority (i.e. there would be no assertion of rights of use over particular numbers).

*Inbound call contains no P-A-ID header field, or an untrusted one*
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. As the 0897 number inserted would be assigned to the gateway network, a conventional certificate could be used.

Outbound international calls
It is recommended that any STIR signalling is passed transparently, other than where ND1016 rule NC2 is invoked to remove a CLI or where the destination network requests otherwise. 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. It should be noted, however, that this could introduce privacy issues as the association of telephone number to originating network could be considered as personal data, and this information is implicit in the STIR signalling: it is a regulatory issue whether this is acceptable.

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

# 8 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 8.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 8.a:** Efficacy of a UK STIR implementation

| **CLI Scenario** | **Interim stages (Phases 1…4)** | | **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** | |
| Residential line (no PN) | NN signed by Originating Network with full attestation | Confirmation that the PN is valid | NN signed by Originating Network with full attestation | Confirmation that the PN is valid | NN signed by Originating Network with full attestation | Confirmation that the PN is valid | NN signed by Originating Network with full attestation | Confirmation that the PN is valid | Assuming SIP, then From header field is populated with the NN (i.e. is the same as the P-A-ID header field) |
| Business line (no PN) | NN signed by Originating Network with full attestation | Confirmation that the PN is valid | NN signed by Originating Network with full attestation | Confirmation that the PN is valid | NN signed by Originating Network with full attestation | Confirmation that the PN is valid | NN signed by Originating Network with full attestation | Confirmation that the PN is valid | Assuming SIP, then From header field is populated with the NN (i.e. is the same as the P-A-ID header field) |
| Business line, static PN (Type 1) | NN signed by Originating Network with full attestation | Confirmation that the PN is valid | NN signed by Originating Network with full attestation | Confirmation that the PN is valid | NN signed by Originating Network with full attestation | Confirmation that the PN is valid | NN signed by Originating Network with full attestation | Confirmation that the PN is valid | |

| CLI Scenario | Interim stages (Phases 1…4) | | 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 | |
| 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 valid | NN signed by Originating Network with full attestation | Confirmation that the PN is valid | NN signed by Originating Network with full attestation | Confirmation that the PN is valid | NN signed by Originating Network with full attestation | Confirmation that the PN is valid | Assumes quality of checking in originating network is sufficient |
| 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 | Originating Network validates number against whitelist, then signs NN with full attestation | Confirmation that the PN is valid | PN signed by the enterprise with full attestation. Originating Network checks this and then signs NN with full attestation | Confirmation that the PN is valid | PN signed by the enterprise with full attestation. Terminating Network checks this | Confirmation that the PN is valid | 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 |

| CLI Scenario | Interim stages (Phases 1…4) | | 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** | |
| 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 | Cannot reliably sign the PN CLI as it doesn't belong to the enterprise |
| 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 | Originating Network validates number against whitelist, then signs NN with full attestation | Confirmation that the PN is valid | PN signed by the enterprise with full attestation. Originating Network checks this and then signs NN with full attestation | Confirmation that the PN is valid | PN signed by the enterprise with full attestation. Terminating Network checks this | Confirmation that the PN is valid | 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 |
| 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 | STIR signing rippled through by diverting network |
| Call Diversion by customer equipment | 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 |
| UK mobile on home network | NN signed by Originating Network with full attestation | Confirmation that the PN is valid | NN signed by Originating Network with full attestation | Confirmation that the PN is valid | NN signed by Originating Network with full attestation | Confirmation that the PN is valid | NN signed by Originating Network with full attestation | Confirmation that the PN is valid | Assuming SIP, then PN (From) is populated with the NN (P-A-ID header field) |

| CLI Scenario | Interim stages (Phases 1…4) | | 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** | |
| UK mobile roaming overseas (home network routing) | NN signed by Home Network with full attestation | Confirmation that the PN is valid | NN signed by Home Network with full attestation | Confirmation that the PN is valid | NN signed by Home Network with full attestation | Confirmation that the PN is valid | NN signed by Home Network with full attestation | Confirmation that the PN is valid | Call routes via home network which behaves as originating network for this purpose – NB this is the default for VoLTE |
| UK mobile roaming overseas (direct routing) | 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? |
| 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 |

| CLI Scenario | Interim stages (Phases 1…4) | | 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 | |
| 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 | Originating Network validates number against whitelist, then signs NN with full attestation | Confirmation that the PN is valid | PN signed by the enterprise with full attestation. Originating Network checks this and then signs NN with full attestation | Confirmation that the PN is valid | PN signed by the enterprise with full attestation. Terminating Network checks this | Confirmation that the PN is valid | 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 |
| 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 | Originating Network validates number against whitelist, then signs NN with full attestation | Confirmation that the PN is valid, 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 valid, but only if the overseas originating network is trusted | PN signed by the enterprise with full attestation. Terminating Network checks this and then signs NN with full attestation | Confirmation that the PN is valid, 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 |

As can be seen from Table 8.a, the interim stage goes a long way to restoring the integrity of CLIs. For calls that are originated in the UK, at the least it allows the originating network which allowed 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. However, what it does not provide is real-time validation that the Presentation Number is legitimate.

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. 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.

# 9 Implications for Network Protocols

## 9.1 UK SIP support for STIR PASSporT

The STIR mechanism (as specified by RFC4474bis) uses a SIP Identity header field containing a STIR PASSporT (as specified in draft-ietf-stir-PASSporT), 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 [7] if/when signing/validation functions are extended to customer/user entities.

## 9.2 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.

## 9.3 PASSporT URI

According to draft-ietf-stir-rfc4474bis-16, 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.

## 9.4 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. RFC4474 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

# 10 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.

## 10.1 Communication Provider Functions

### 10.1.1 Call Server

The main change to Call Server 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.

### 10.1.2 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.

### 10.1.3 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.

The text in Section 7 makes an assumption that a certificate which is used to generate keys would accommodate many numbers, but it is as yet unclear whether and how this will be possible, or alternatively whether only a limited volume of numbers could be associated with each certificate. Either way, each operator will be dealing with many certificates and associated keys, and in turn these certificates will need to be authoritative for large volumes of numbers, which may not be contiguous due to number portability factors. These issues are under discussion at the IETF at the time of writing of this report, and it is possible that there could be scalability problems.

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.

## 10.2 Central Functions – Certificate Authority

The Certificate Authority functionality will need to be implemented by a central agency, for example appointed by Ofcom or UK communication providers 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 mapping of which communication provider (or even enterprise in the Final Stage solutions) has the rights to use which telephone number. Due to number portability, the only way to accomplish this is to have a database of which individual number is assigned or ported to each communication provider. The cost of this should not be under-estimated; it will be a large database (of the order of half a billion numbers, depending on whether all numbers allocated to communications providers are populated, or just live ones). Also, initial population could be complicated by data integrity issues that have developed over 20+ years of number portability. This said, some of the cost could be mitigated by using any database created for applications wider than STIR itself (for example number portability).

The Certificate Authority as envisaged in Section 7 of this report would not need to be queried in real-time, and as such arguably may not need 99.999% resilience. However, the loss of connectivity would have ramifications for STIR operation. On the originating side, any loss of the function would mean that numbers being provisioned (whether new provides or ports) would not be associated with a certificate, hence would not have their outbound calls signed. Similarly, at the terminating end loss of functionality would mean that newly renewed certificates could not be referenced, thus meaning that the ability to fully check signatures could be lost (the ability to see that calls are signed would not be lost, but the ability to check the signature would be).

# 11 Conclusions and Next Steps

This document has set out what would be achieved by a staged implementation of STIR technologies in UK networks. The Interim Stage would in principle provide surety of which originating CP admitted a call into the public network, and 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.

Further, the solution as set out in this document would require a database of which numbers are assigned for usage on which networks, i.e. a central 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.

It is possible to launch STIR without a numbering database, but this would only provide a pointer back to the network that originated a call, rather than whether they had any rights to use the associated CLI. Given the large volume of communications providers in the UK, this would be of limited value.

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.

In developing technical standards, NICC will ensure that implementation of STIR is not precluded. However, it is considered premature to develop specific UK standards for STIR implementation. Should a suitable business case be proven for UK STIR, NICC will develop the associated technical standards as UK networks evolve to the point 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 |
|---|---|---|---|---|---|
| 1 | Start up. Voluntary signing. | Creation of architecture, including numbering database | NN by originating network.<br>• For PN being Type 1 & 2, Full Attestation.<br>• For PN being Types 3-5. Partial Attestation<br>• For inbound international, Gateway Attestation | N/A | Gets the ball rolling |
| 2 | 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 |
| 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 |

| 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 three 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 draft-ietf-stir-rfc4474bis 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 draft-ietf-stir-rfc4474bis 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.

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 draft-ietf-stir-rfc4474bis 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.

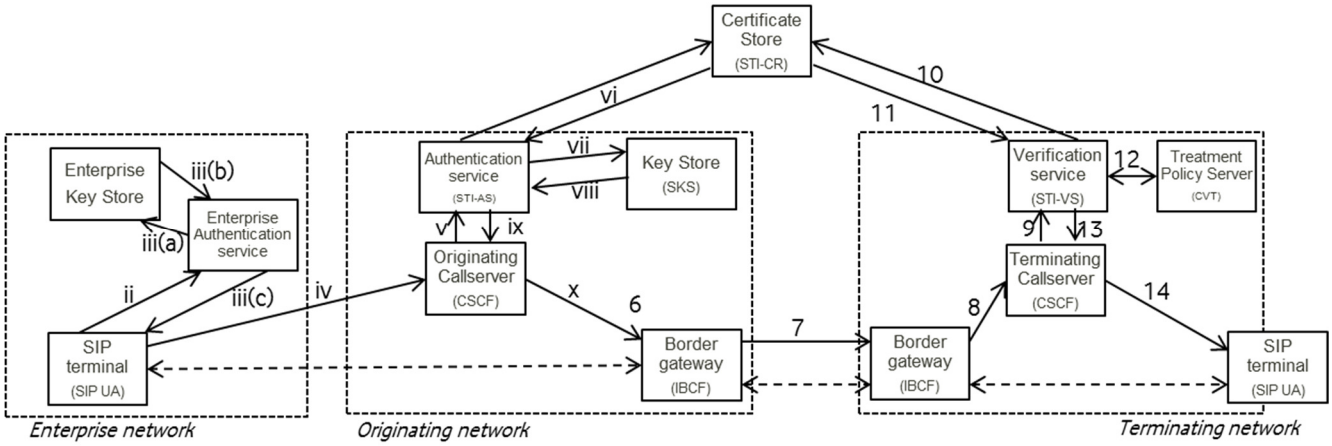


**Figure B.1:** Final Stage Variant II

Variant III – originating customer signs PN, Terminating 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. 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.

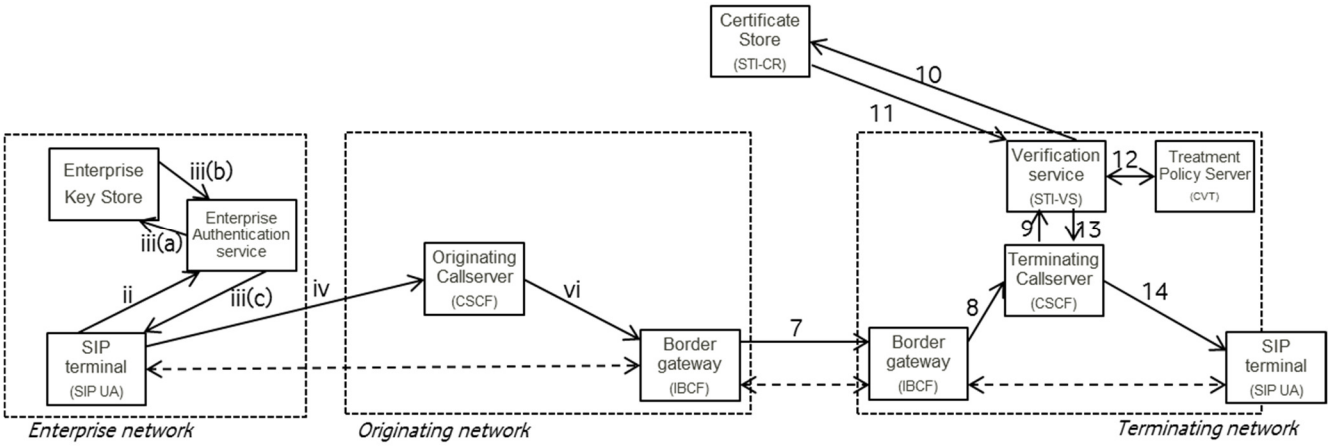


**Figure B.2:** Final Stage Variant III

# 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 5 best meets the identified needs;

- 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 5 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.

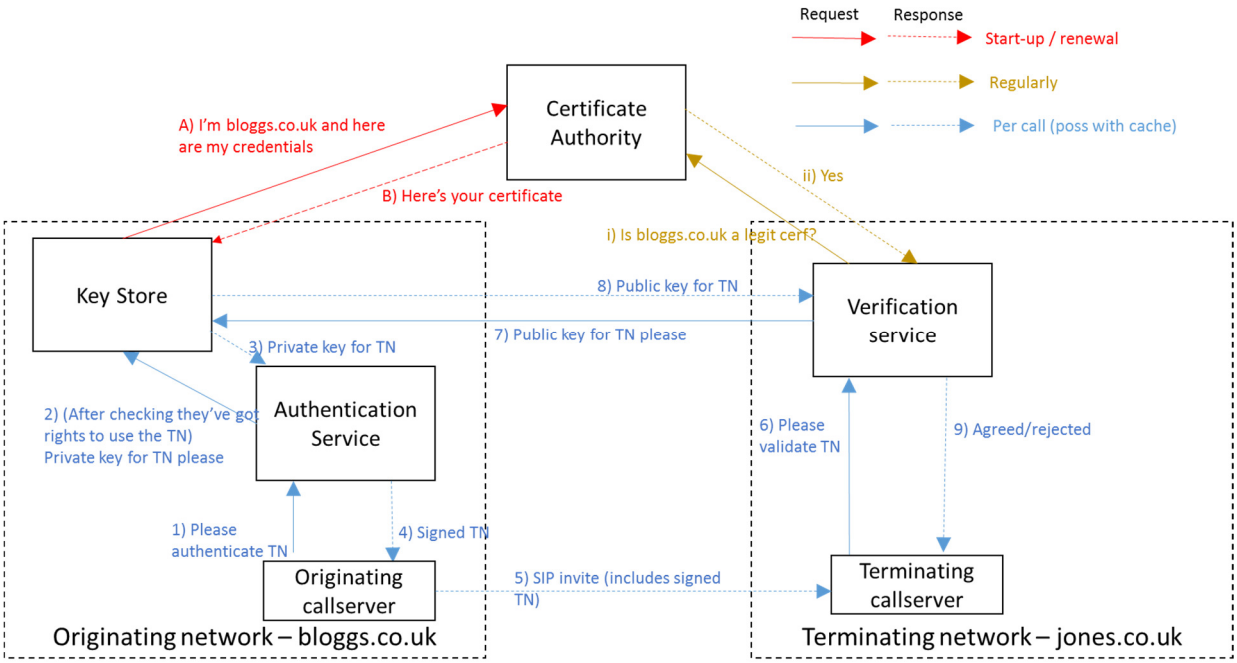


**Figure C.1:** Approach 1

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. However, STIR as considered in this report requires end-to-end SIP, which will not be universally available in the UK for perhaps a decade: as such there is the opportunity to implement a CDB prior to STIR implementation, hence negating the benefit of using Approach 1 as a launch phase. For these reasons, it was concluded that Approach 1 is not NICC's preferred option.

## 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.

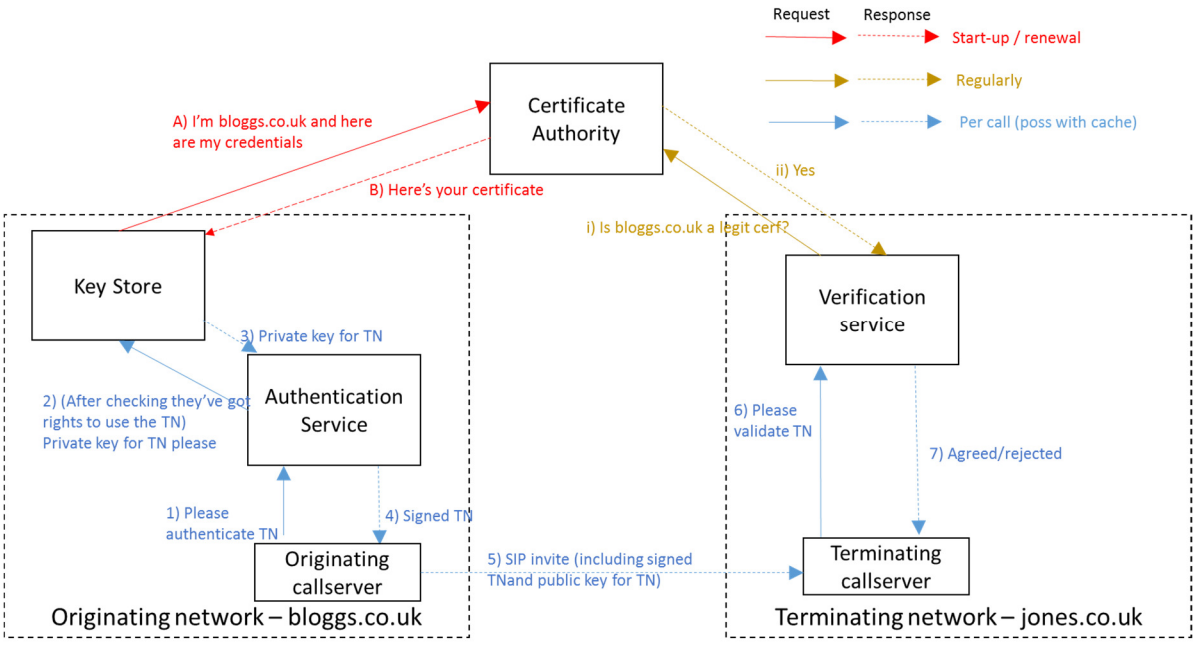


**Figure C.2:** Approach 1a

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.

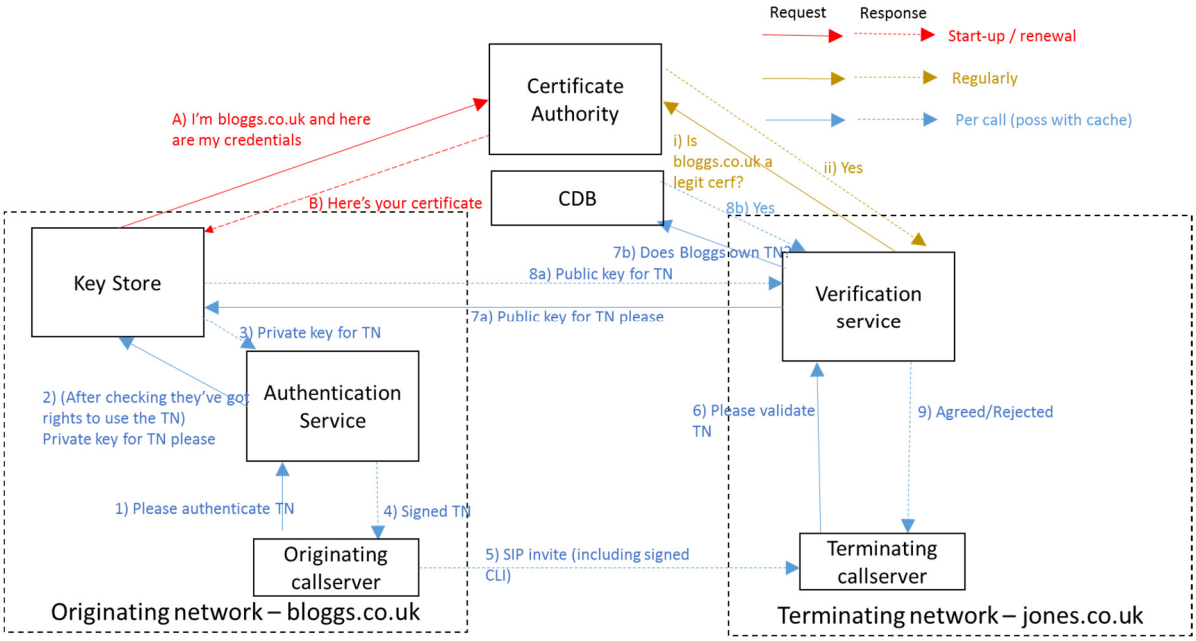


**Figure C.3:** Approach 2

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. However, set against this it either requires two sequential queries at the Verification Service – with associated potential post-dial delay – or requires the Verification Service to be able to control two parallel queries and join the results. Therefore, although Approach 2 was considered a workable option, it wasn't considered the preferred one.

## 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.

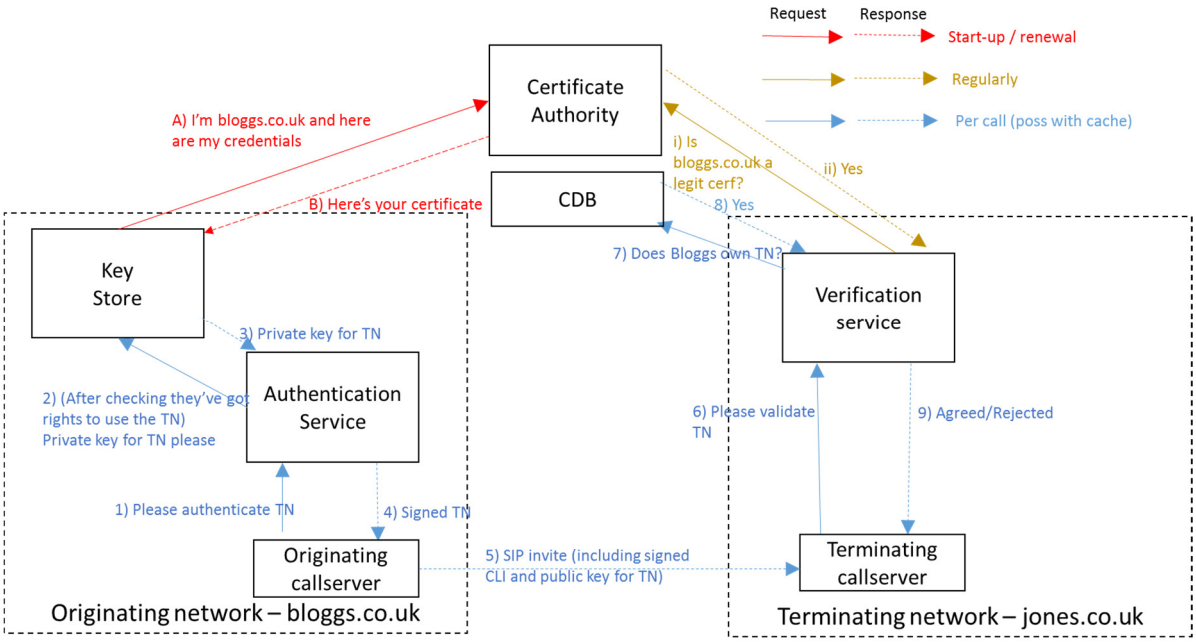


**Figure C.4:** Approach 2a

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.

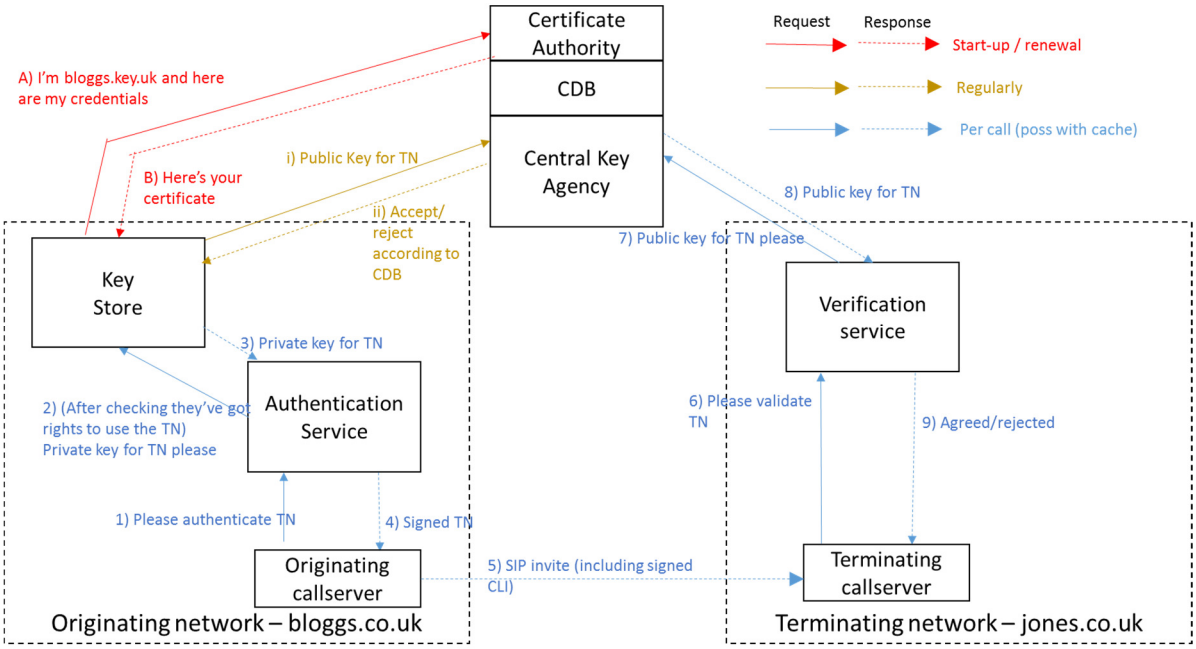


**Figure C.5:** Approach 3

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.

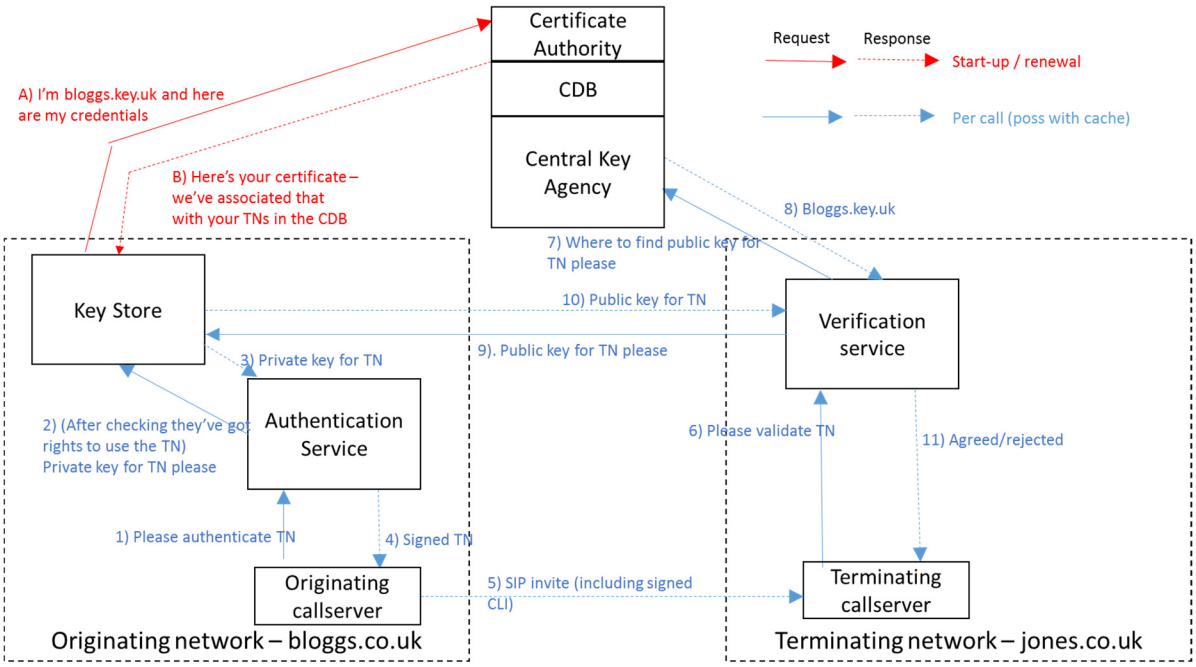


**Figure C.6:** Approach 4

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.

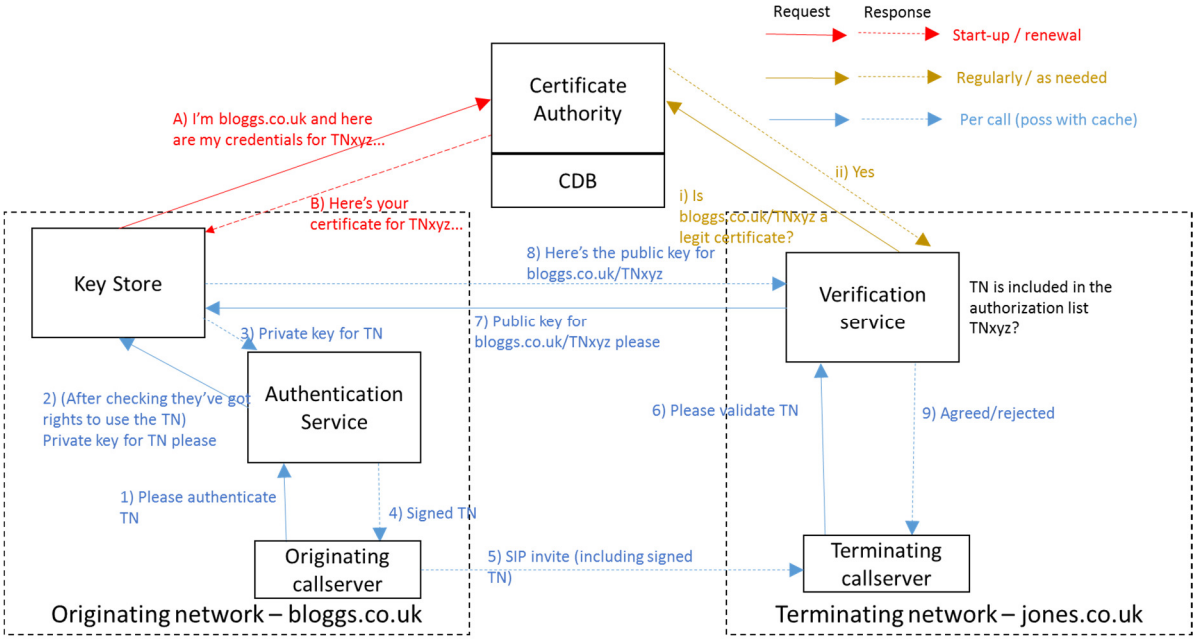


**Figure C.7:** Approach 5

With proper management of the mapping of telephone numbers to certificates/keys, 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 B.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.

Finally, although it is difficult to predict international developments, it is believed that if/when SHAKEN evolves to a solution where the originating networks' right to use numbers is checked, a model similar to this approach could be adopted. For these reasons, NICC recommends this be the preferred UK solution for STIR implementation.

## 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.

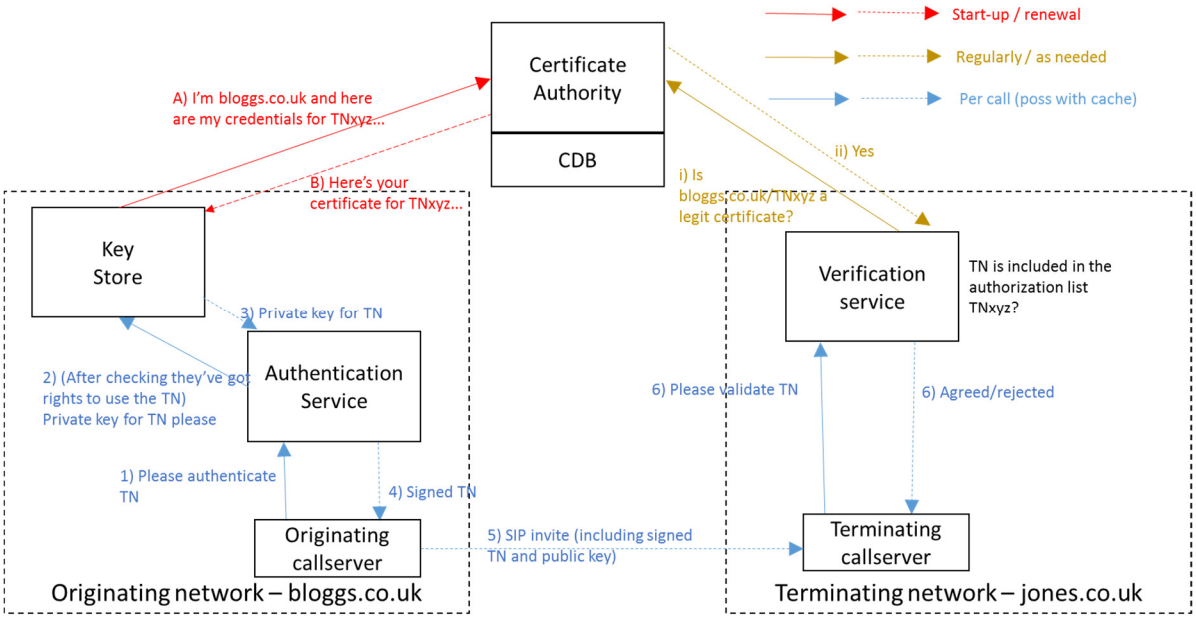


**Figure C.8:** Approach 5a

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.

# Bibliography


- ATIS-1000080 (11th July 2017): "Joint ATIS/SIP Forum Standard – Signature-based Handling of Asserted information using toKENs (SHAKEN): Governance Model and Certificate Management".
- ATIS-0300116 (5th December 2016): "Interoperability Standards between Next Generation Networks (NGN) for Signature-Based Handling of Asserted information Using Tokens (SHAKEN)"
- IETF RFC7340: "Secure Telephone Identity Problem Statement and Requirements"
- draft-ietf-stir-oob-00 (3rd July 2017): "STIR Out of Band Architecture and Use Cases"
- draft-ietf-stir-certificates-14 (9th May 2017): "Secure Telephone Identity Credentials: Certificates"
- draft-ietf-stir-passport-11 (9th February 2017): "Personal Assertion Token (PASSporT)"
- 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 |
| | | |