

Functional Requirements for an architecture to support Enhanced Communications Services

[TERENA Task Force Enhanced Communication Services](#)

Deliverable 'Requirements for the Architecture'

Version 1.0, May 2007

Erik Dobbelsteijn (SURFnet)

Contributors: Cătălin Meiroșu (TERENA), Dimitris Daskopoulos (GRNET), Fabio Vena (SWITCH), Jan Ruzicka (CESnet), Kewin Stoeckigt (AARnet), Marco Sommani (GARR), Olav Kvittem (UNINETT)

This document describes the functional requirements of an architecture that intends to connect real-time Enhanced Communication Services (ECS) domains in a secure and scalable way. It reflects the result of discussions amongst National Research Networks and will be used as a starting point for the design of an infrastructure that stimulates the use of ECS while protecting users from malicious use from unidentified originating sites.

Objective

Currently, academic institutions are in the process of migrating to IP telephony solutions. These new systems are voice centric, and are usually referred to as 'IP Telephony' solutions.

At the same time, some institutions are implementing an infrastructure that also supports video, instant messaging, presence, and possibly even desktop sharing. This ensemble of various media, sometimes referred to as 'rich media' and further referred to as 'Enhanced Communication Services' or 'ECS', is the target of the TERENA Task Force 'Enhanced Communication Services.' Such an infrastructure can be based on the SIP.edu open-source model or commercial software and hardware platforms, including client software for all end-users.

Implementing ECS allows for widespread communication within an institution, but also inter-institutional. The ECS implementation provides end-users with an account that they can use to set up multimedia sessions with end-users within the same domain or other domains. Each implementation is in itself an 'island'.

The current mechanisms to make the island reachable are very open and unregulated. This situation is comparable to the way that domains are opened for e-mail, which makes it easy for spammers to also reach all users in the domain. It is very likely that 'rich media spam' will also target non-suspecting users by employing real-time media. As opposed to e-mail, this form of spam is much more intruding, because it is not easy to be identified as such by any sort of spam filter. A call might have to be answered before it can be recognised as being unwanted.

Hence the need for an architecture that connects the various infrastructures, that connects the islands of real time communication, while protecting the end-users from unwanted interaction

Starting points

The requirements are a result of a couple of high-level starting points that were defined by the participating NRENs and institutions in the TF meetings. The main starting points are:

- Build a coherent infrastructure, connecting institutions, NRENs and telephony providers (all in fact different domains or 'islands')
- Avoiding the creation of an island of islands
- Ensuring the trustworthiness of the identity of callers within the academic community
- SPIT handling
- Use of Open standards
- Support of a wide variety of vendor software and equipment
- Scalability of the architecture
- Ease of use for the end-user

We foresee cases where ECS end systems will have to be able to present the end-user with means to determine if the identity of the caller can be trusted. It therefore is desirable to define and test a set of rules, aimed at setting up a federation enabling SIP communications between mutually trusted partners, at least inside the NREN community. The rules should not put restrictions on the variety of usable end systems. The rules must be enforceable at least on the open source based ECS systems commonly used in the NREN community (which are for SIP: SER, OpenSER, Asterisk).

Elements and Relationships to be supported

The architecture that will result from these requirements will define objects that will implement well-defined functionalities, and the relationships between those objects.

Elements

Within the architecture, the main elements are:

- The end-users that have a communication need which is met by their home institution
- The home institution or campus, providing or planning two types of services:
 - PBX-style telephony services, either based on traditional TDM based PBXes or IP based PBXes.
 - Enhanced Communication Services including video, Instant Messaging
- The NREN, possibly providing means to connect campuses for both Telephony and ECS
- Telecom providers that can handle (telephony) sessions to and from the PSTN

Relationships

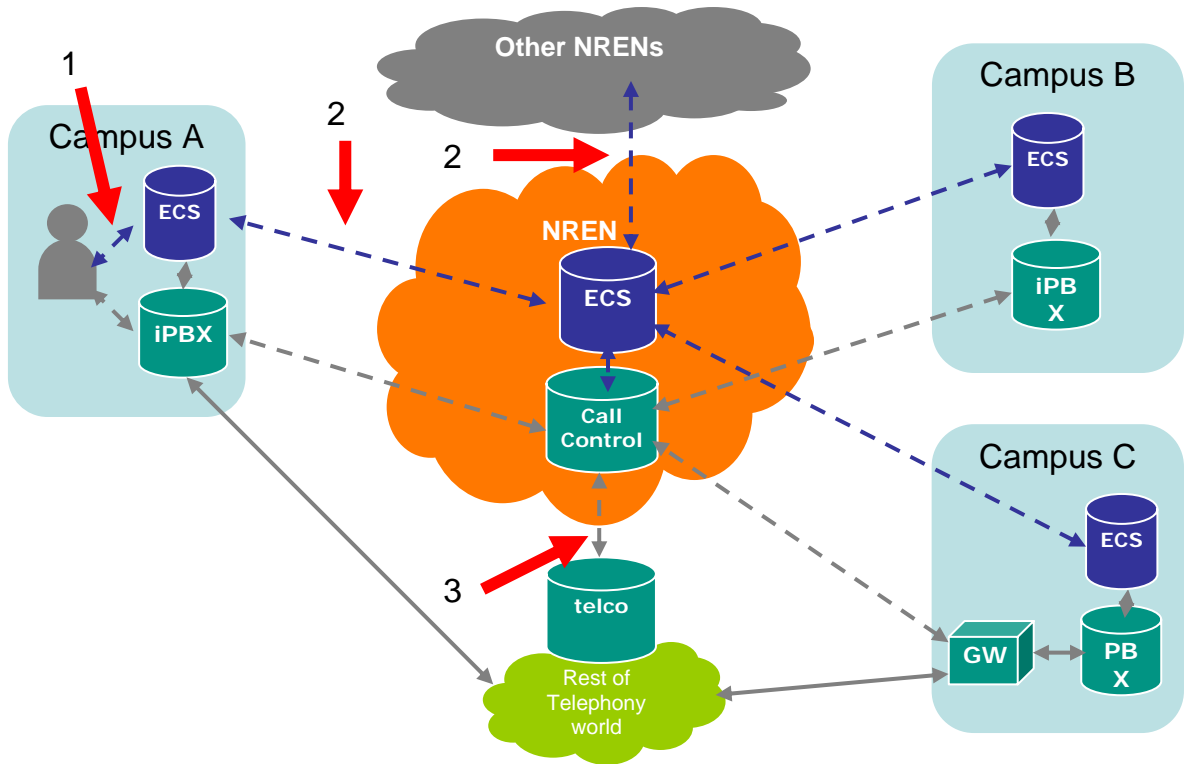
The relationships that have the primary focus in the architecture are those between:

1. End-users and their home ECS domains
2. ECS domains within the NREN community (including other NRENs than those of their home domain)
3. ECS domains and commercial parties (Telco's, IM Service Providers)
4. This constitution of domains and other federations.

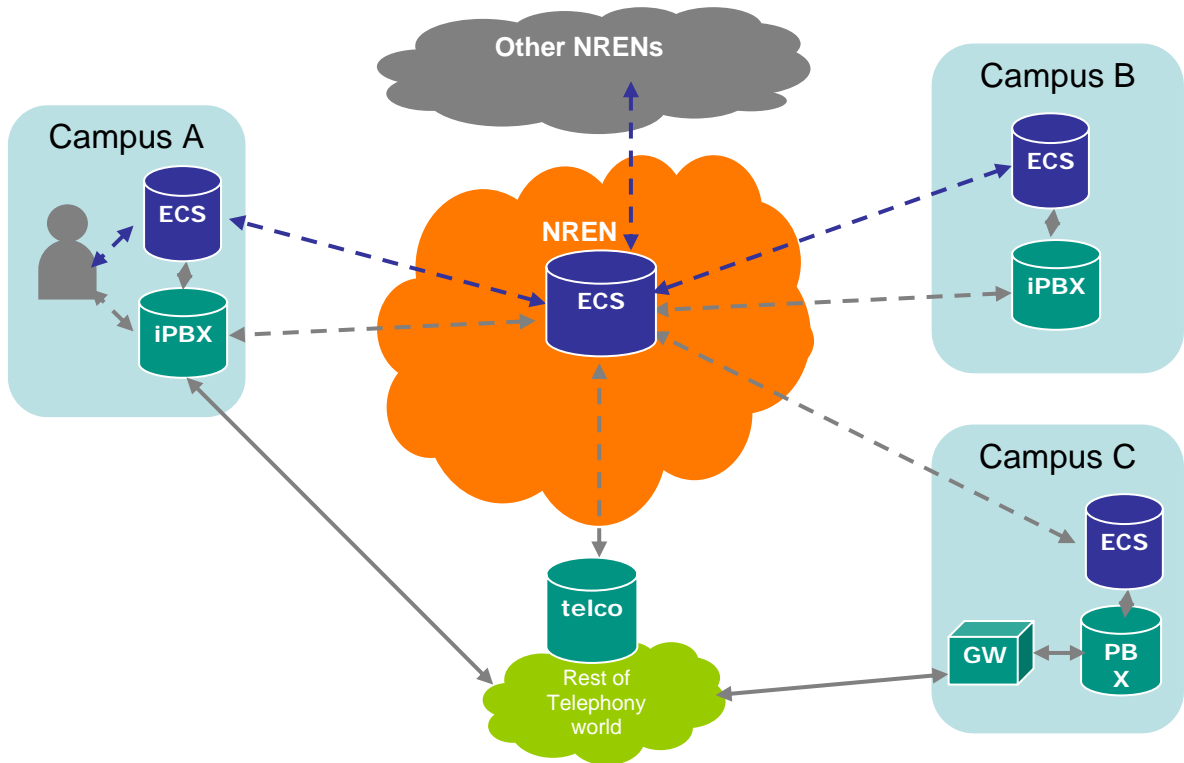
These relationships are depicted in the picture below. The two parallel infrastructures in the diagram are representative for the current state of implementations.

Institutions tend to have a telephony-only focus with respect to functionality of their real-time infrastructure, while ECS functionality is usually in a pilot phase and is built in parallel but separately from the telephony infrastructure. The level of integration between the parallel infrastructures is not present or limited.

Oddly enough, in each of the two infrastructures, contradictory developments take place. On the one hand, iPBXes hold the promise of offering SIP, but enable their users rarely to be reached by SIP from the IP world. The connection with the rest of the world is still a PSTN connection, and because of the PSTN roots of the manufacturers and the interests of suppliers which are often also PSTN carriers, these 'islands' are not likely to open up easily. On the other hand, ECS infrastructures are open by default, especially if they are based on SIP.edu. Administrators managing such infrastructures are considering SPIT prevention, which might lead to narrowing down the reachability of their implementation.



Connected elements in this example form a 'trust fabric', establishing a federation between domains. This hierarchical structure is not the only possible way to connect different domains, so this picture should only be interpreted as a diagram of trust relations. More on this is described in the 'Requirements' section. Our target architecture of course combines both IP Telephony and ECS infrastructures:



Roles of the elements

End-user terminal

This is the agent –either dedicated hardware or software on a generic device- that interacts with the user. It takes care of the user interaction, connectivity, media conversion and signalling.

Lookup mechanism

A lookup mechanism provides translation from numbers or URI's to the routing entities that have routes to the called end-user terminal.
(compare: GDS, ENUM, SPIRIT)

Routing

Routing entities take care of passing session requests for a number or URI to the terminal of the called end-user. Routing is possible between end-user terminals, institutions, NRENs and commercial service providers.
The routing setup should be scalable: domains complying to the same set of agreements corresponding to this architecture can delegate the interconnection to 'domain interconnect entities'.

Supported scenarios

The following session initiation scenarios will be covered by the architecture. Each scenario can also be initiated vice versa.

1. Set up of a voice session between two end-users on two different campuses
2. Set up of a voice session between an end-user and a PSTN number
3. Set up of a voice and video session between two end-users on two different campuses
4. Set up of a session between two end-users
5. Invite more end-users into an instant messaging session

Supported Media

Critical in the interoperability between ECS islands is the use of the same type of media transport. This chapter provides an overview of the set of minimal requirements:

audio

Two-way audio should be based on at least the G.711u and a-law codec.

video

Two-way video should be based on at least the H.263 codec with a CIF resolution and a frame rate equal to or higher than 25fps.

presence

End-users can define who their 'buddies' are, and show their status (available, away, busy, offline)
optional: end-users can show their geographical presence, apart from their status.

instant messaging

End-users can send text messages of 256 characters long to another end-user, provided that they are 'buddies'. Multiple end-users can share consecutive text messages, each 256 characters long (also refers to as 'chat').

Screen sharing (optional)

parts of or an entire computer screen can be shown to other end-user(s) in a session, up to a resolution of 1024x768 pixels and a refresh rate of 5 frames per second

Application sharing (optional)

Other end-users can manipulate the shared application of an end-user with their mouse and keyboard, after the application owner allows this.

Functional Requirements of the infrastructure

The architecture should support a number of features to meet the general starting points:

1. Verification of the identity of the domain of the session initiator
2. Verification of the identity of the session initiator
3. Notification of trustworthiness of Session Initiator ID
4. Lookup mechanism to find users if they want to be found
5. Support for numeric-based dialling schemes (comparable to GDS)
6. Support for URI-based dialling schemes
7. Routing
8. Signalling conversion/interop
9. Media conversion
10. Forking (& breaking down) of sessions
11. Authentication of the end-user at his home institutional ECS infrastructure
12. Encryption of session signalling and media between the end-users
13. Provision of hooks to support Lawful interception
14. Support for Traffic troubleshooting

More functional requirements are worked out in more detail below.

Numbering

End-users should be reachable both by URI and E.164 number. For end-users that are supplied with an ECS service which does not offer PSTN calling, an alternative dialing scheme like ISN is an option.

Firewall traversal and NAT traversal

For both signalling and media it should be possible to traverse firewalls and NAT based routers.

Sessions from and to a group of end-user terminals behind the same firewall can be proxied (for example using an SBC or B2BUA), or the firewall has to be aware of sessions and pass them through.

Encryption

Both signalling and media should be able to be encrypted, optionally. Minimal requirement is encryption by AES 128 bit.

'voice'mail

Recording of an audio message by an answering-machine-style service when a called end-user is not available should be possible, which the user can later retrieve through a voice connection (required) or voicemail-in-e-mail (optional). The option to also record and retrieve video is optional

multipoint sessions

It should be possible for two or more end-users to create a multi-party session, for at least the following media:

- audio
- video
- screen sharing
- instant messaging

Connections to the PSTN

Dependent of the role of the NREN, the national architecture can support routing calls to and from the PSTN by itself, or it makes it possible for telephony providers to handle the PSTN traffic.

Institutions can make direct telephony peerings with telephony providers using the IP or Ethernet connectivity that the NREN offers. Note that the word 'peering' is being used differently in the telephony domain as it is in the internet domain. Another option is to make use of ECS routing nodes within the NREN. The protocol(s) used to connect the NREN infrastructure with the telephony providers should be based on open standards and thorough agreements on how to deploy the standard should be made regarding the number plan, routing, parameters in the protocols, billing, and regulatory obligations.

SPIT and SPIM handling

While ECS domains will hopefully still accept sessions in an open fashion, even when the risk of SPAM for Internet Telephony and for Instant Messaging increases, an initial step is taken to provide the user with means to recognise and ignore unwanted calls or messages.

Simple measures like black lists (on a per-user basis and on an infrastructure level) can offer a crude way to ignore sessions from unwanted domains. It would be preferable to present the end-user with a 'trust indication', showing if the identity or at least the domain of the session initiator can be trusted. The end-user should be in full control of which calls to receive, which calls not to receive and how they will be tagged.

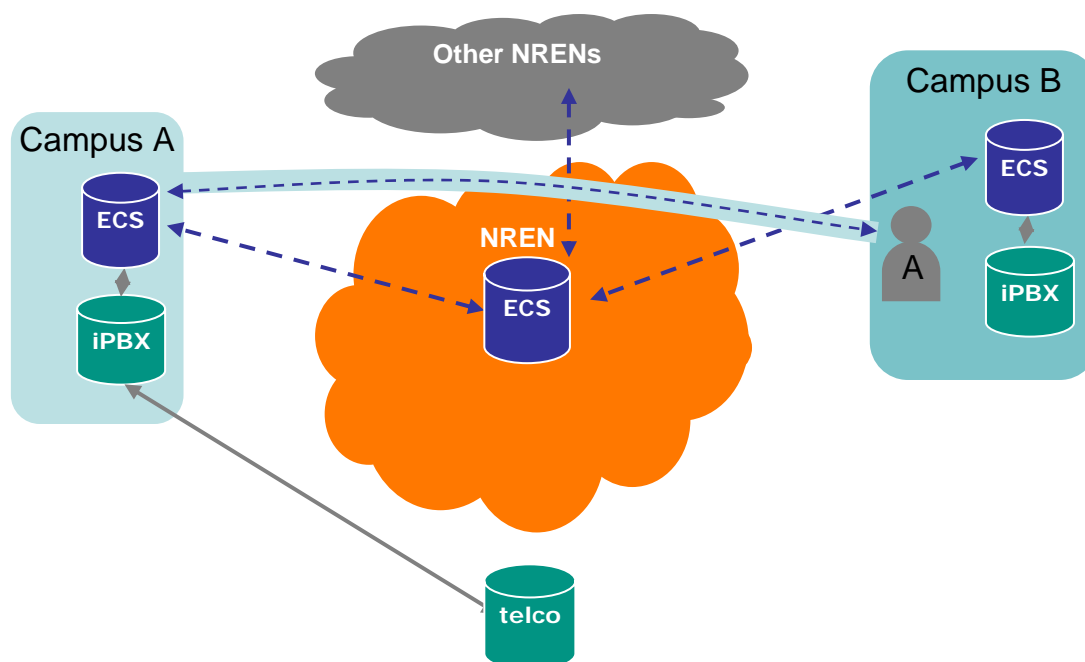
One way of trying to establish the level of trust of an incoming session is to enable domains to reach each other in a trusted way. This may imply a federation-like approach towards interconnecting domains. This federation relates to the back-end (infrastructure interconnection) of the domains. The domains in the infrastructure can accept each others sessions based on a trust that can be established in various ways. Static 'peerings' are a start, but the method does not scale when many domains are involved. A hierarchy of domains where each NREN provides for the top-level trust node for its country and international nodes connect these national top-level nodes corresponds with mechanisms like the Global Dialling Scheme and eduroam. It takes a lot of work to establish all the relationships in the hierarchy. It should be possible to dynamically trust domains based on an existing common trust relationship that for instance resides within PKI or DNSsec, or a trusted 'lookup' node.

A totally different approach is an end-to-end approach (whereas the previous is a hop-by-hop approach). End-users can directly exchange identity information during session setup which is verified by nodes that are not necessarily part of the real-time infrastructure.

Roaming

In all cases, the end-user must log onto his particular home domain infrastructure. This is also true when the user is roaming on a different network. Like eduroam shows, once a user has obtained access to a network (through a federated log on mechanism, like eduroam), a user should log on to his home applications, either by directly contacting application servers using end-to-end security or through a VPN tunnel. The real-time performance of VPNs may impose problems for services involving media exchange and should be tuned carefully.

This mechanism is shown below.



In special cases, it is possible to federate the actual authentication of the end-user. For instance, when the NREN provides an end-user service to users in different domains, the logon can be distributed to the different home domains.

Lawful interception

Currently, lawful interception requirements differ per country, are often times not clear yet and the question remains whether NRENs fall in the category of parties obligated to implement it. This makes it impossible to define requirements in this document regarding this issue.

Agreements and Procedures

The introduction of a new domain into the trust fabric should be organised in such a way that those responsible on both sides of the connection can be identified properly and are proven to be allowed to act on behalf of the institution or NREN they represent.

An Acceptable Use Policy (AUP) for end-users should be formulated to cover responsibility for misbehaviour of end-users.

Agreements between institutions and other institutions or NRENs should delegate the responsibility of tracing misbehaviour and take measures against the end-user or the appropriate party.

Glossary

The terms and abbreviations used in this document are described in the 'glossary' section of the TERENA IP Telephony Cookbook:

<http://www.terena.org/activities/iptel/chapters/Glossary.pdf>