Enabler API IMS

Technology Reports

SEN Infrastructure for Configuring a Network Cloud

In today's business environment, communications operators need to develop and provide diverse, value-added services and bring them to market promptly. To meet this need, we have developed an infrastructure for achieving a service enabler network that combines multiple enablers in a network cloud. This approach makes it possible to provide a service by developing only a service scenario. It also shortens the service development period so that a variety of communication services can be provided in a timely manner.

Core Network Development Department

Subrata Biswas
Yoko Iimura
Yu Kojou
Karin Sugiyama
Takeshi Yamamoto†

1. Introduction

With the dramatic proliferation of smartphones and the rapid rise of global Internet players in recent years, communications operators are running the risk of becoming "dumb pipes." To prevent this from happening, it is becoming increasingly important for them to come up with ways of providing value-added services. At present, however, the mechanism for providing services is complicated, and individual services in which common functions and service-dependent functions are closely intertwined must be separately optimized. Much time and cost are therefore required before the service can be provided.

In response to this issue, we have developed an infrastructure for realizing a Service Enabler Network (SEN) infrastructure with the aim of achieving value-added network services quickly and cost effectively as now demanded of communications operators. The SEN infrastructure is one element for configuring a "network cloud" as a basis for providing added value through the advanced information- and communication-processing capabilities of the network. It enables multiple enablers to be combined so that a service can be provided quickly by just developing a service scenario. This article describes the configuration of the SEN infrastructure and presents examples of services achieved through its application.

2. Infrastructure Configuration

The overall logical configuration of the SEN infrastructure is shown in **Figure 1**. The SEN infrastructure consists of (1) modularized service functions (enablers) that can be used to provide a variety of services and (2) common control functions for combining different enablers and achieving a service scenario. This infrastructure has two key features:

- Shortens the service development period by making flexible use of modularized service functions and common control functions while minimizing the impact on equipment, localizing the scope of that impact and optimizing the scope of

† Currently Services Platform Department

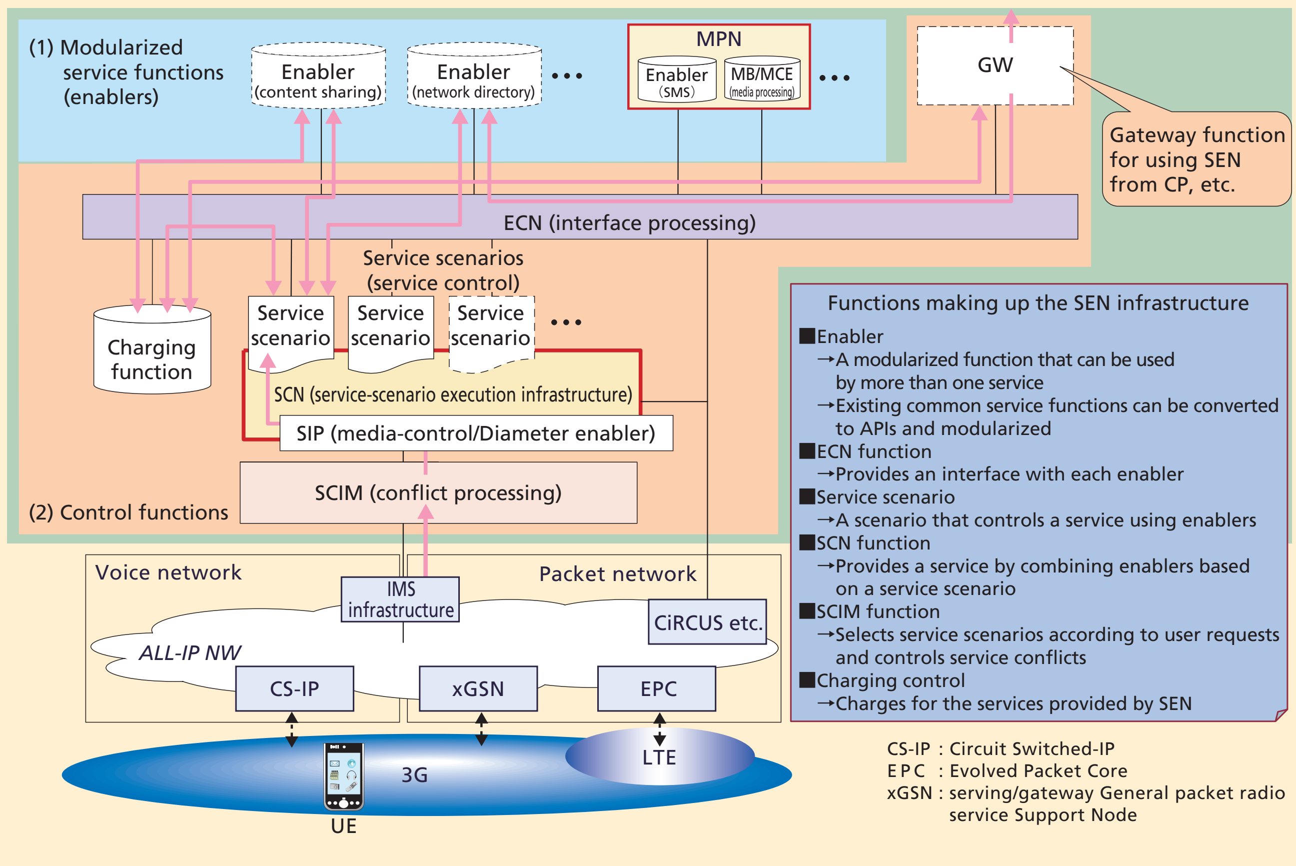

Figure 1 Overall logical configuration of the SEN infrastructure

testing.

- Constructs a mechanism that enables a variety of enablers to be used on the basis of a service scenario without having to be concerned about functional differences between those enablers.

2.1 Essential Functions

The aim of developing a SEN infrastructure was to "develop a network equipped with basic functions, that is, telecom functions, enabler functions, media control functions and Web access functions." In this way, it would be possible to provide services simply on the basis of developing service scenarios, whether those services are ones using previously implemented enablers or ones not requiring special enablers.

First, with respect to telecom functions, we equipped the SEN infrastructure with functions to support the call models needed when providing new services in addition to the basic call model provided by Session Initiation Protocol (SIP)[*1] Here, so as not to adversely affect IP Multimedia Subsystem (IMS)[*2] equipment, we added a SEN infrastructure-linking function to IMS equipment to link existing services with services on the SEN infrastructure. We also defined a service-conflict processing function in the form of a Service Capability Interaction Manager (SCIM) that can be used in common by service scenarios and installed this function in the Service Composition Node (SCN)[*3] [1].

Next, as for enabler functions, we have been converting existing functions such as Short Message Service (SMS)

---

*1 **SIP**: A call control protocol defined by the Internet Engineering Task Force (IETF) and used for IP telephony with VoIP, etc.

*2 **IMS**: A call control procedure that realizes multimedia communications by consolidating 3GPP standardized communication services offered over fixed and mobile networks through the use of SIP, which is a protocol used on the Internet and in Internet phones.

*3 **SCN**: A node that combines a collection of enablers to execute a service scenario. Includes the SCIM function.

and location information into application programming interfaces (APIs)*4 so that they can be used as enablers. We have also developed an Enabler Connection Node (ECN) as enabler accommodation equipment to simplify the introduction of new enablers.

To provide media-control and Web-access functions, we use the previously implemented SCN and Media Processing Node (MPN)*5 [2].

2.2 Function Overview

In this development, we implemented a SEN infrastructure-linking function in the IMS infrastructure and developed SCIM as a new control function in the SEN infrastructure. We also introduced ECN to minimize the amount of network-specific customization work when introducing a new enabler. These functions are described below.

1) SEN Linking in the IMS Infrastructure

Linking between the IMS and SEN infrastructures is shown in **Figure 2**.

A connection between IMS and SEN is made from the IMS side, which treats the SEN infrastructure as a single Application Server (AS)*6. For this reason, a Serving Call Session Control Function (S-CSCF)*7 in IMS conforming to 3GPP TS23.218 and 3GPP TS29.229 specifications references Shared initial Filter Criteria (SiFC)*8 so as to determine the type of AS to connect to and the procedure required for connection. This makes it possible to sort through a number of services and initiate the proper sequence. However, SiFC is restricted to static logic, that is, it is defined in such a way that service and session*9 states are not passed between ASs. To eliminate this restriction, we made it possible in this development for an existing AS to be aware of service launchings. Specifically, we added service control information to the response signal when S-CSCF makes a connection with a new AS and enabled SCIM to perform conflict control between existing services of existing ASs and new services of the SEN infrastructure (new ASs). This enables conflicts between services executed by existing ASs and new services executed by new ASs to be assessed thereby enabling services to be combined within the same session.

The following outlines the process

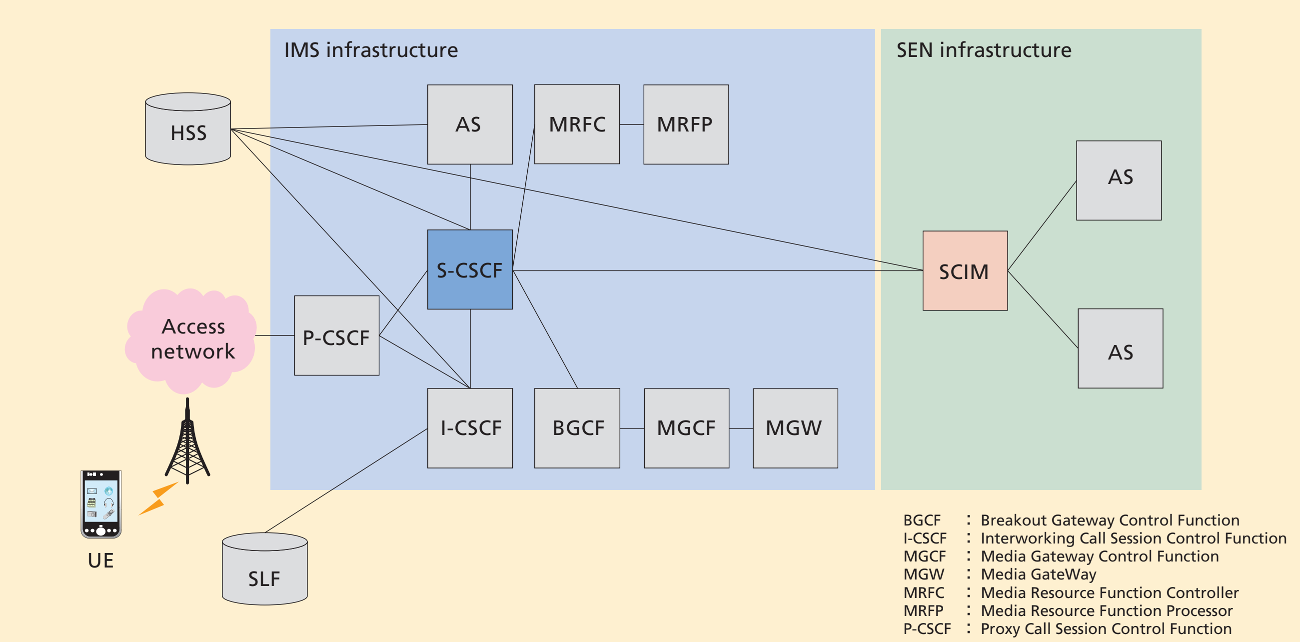

Figure 2 Linking between IMS and SEN infrastructures

*4 **API**: An interface that enables a function provided by an enabler to be used by other equipment.

*5 **MPN**: A node of the NTT DOCOMO core network. It provides state for various media services including voice answering, melody call and other voice media services; and video media services such as videophone answering.

*6 **AS**: A server that executes an application to provide a service.

*7 **S-CSCF**: A SIP server performing terminal session control and user authentication.

*8 **SiFC**: Criteria for determining which AS to send a request signal or other message to and the function for doing so.

*9 **Session**: A series of communications exchanged between a client and a server or between servers.

flow for providing services by existing ASs and by new ASs in the same session. When adding a new service to the SEN infrastructure, the Home Subscriber Server (HSS)[*10] sets new information for connecting with SCIM in SiFC. Then, at the time of a call connection, S-CSCF references user-profile information in SiFC and prioritizes the selection of SCIM. Now, on receiving a signal from S-CSCF, SCIM determines whether a SEN service needs to be launched, and if so, adds a "SEN connection request" to the response signal sent to S-CSCF. Subsequently, whenever S-CSCF sends a connection request to an existing AS based on SiFC information, it will include "SEN connection request" information in that connection-request signal. This enables the existing AS to be aware of services running on the SEN infrastructure.

2) SCIM

The role of SCIM in the provision of additional services using the SEN infrastructure is as follows. On receiving a connection request from S-CSCF, SCIM determines the origination and termination of the call, gets the user profile, and based on that user profile and service configuration data, launches a service scenario at the appropriate time for the user in question. A major advantage of introducing SCIM is that the provision of new services and changes to conflict conditions in existing services can be supported by simply updating various configuration data without having to update actual files. Here, however, as it is unrealistic to introduce conflict control that covers every possible pattern, it was decided to establish conditions for the types of conflict control to be supported and to enable conflict control to be changed within that range.

The following policies have established for assessing service conflicts when implementing the SEN infrastructure (**Figure 3**).

- SCIM determines which service from among existing-AS services and new-AS Services to launch without existing ASs having to be aware of individual services running under new ASs.
- Given an existing-AS service and new-AS service that are associated with the call session and that use the same service-provision timing, either one or the other will be provided.
- If service-provision timing is different, or if it is the same but with either of the services being independent and not associated with the call-connection session, the existing-AS service and new-AS service may be used simultaneously.

3) SCN and MB

The configuration of SCN and Media Broker (MB)[*11] is shown in **Figure 4**.

In addition to a service scenario section (SCN-AS), SCN consists of an enabler section, which provides the service scenario with abstract APIs for call control and media processing, and a function for storing service-specific user data (SCN-DB).

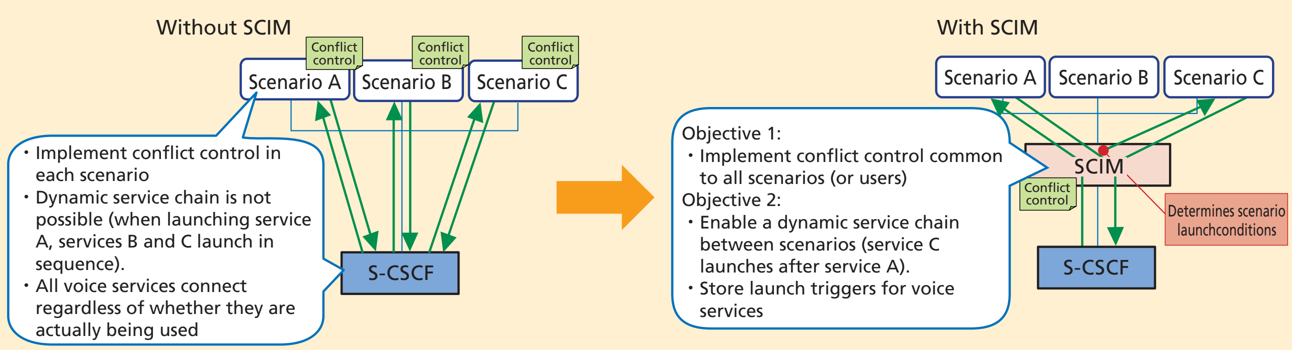

Figure 3 Conflict control before and after introducing SCIM

*10 **HSS**: The subscriber information database in 3GPP mobile communication networks. Manages authentication and location information.

*11 **MB**: A node that acts as an intermediate processor of text, video and other forms of media. It is in charge of protocol and media conversion, data duplication, etc. for one or more connection destinations.

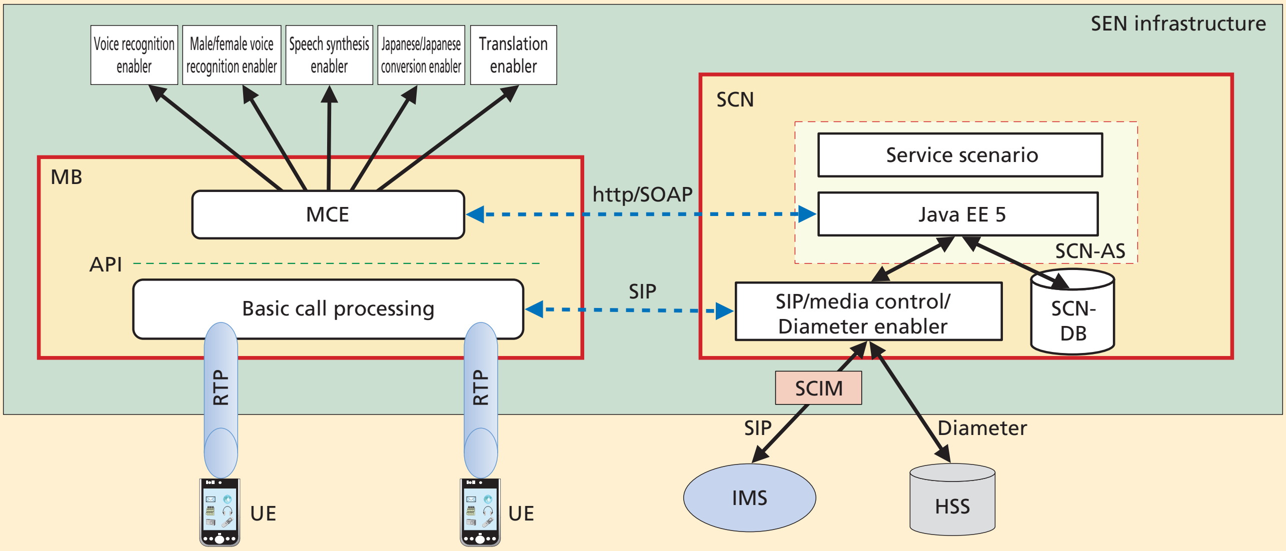

Figure 4 Configuration of SCN and MB

- SCN

SCN-AS is an application program running on a Java® EE5[*12] server. It operates as a Web server for handling Web access from users, and at the same time, it communicates with enablers having a Web-based interface both inside and outside of SCN.

The enabler section is equipped with a call-control enabler that abstracts SIP interworking[*13] with IMS, a media-control interface that abstracts connection processing with media-processing equipment, and a Diameter enabler (3GPP TS 29.328) that abstracts Diameter[*14] signal processing with HSS [1].

SCN-DB manages service-specific user profiles by interworking with SCN-AS.

- MB

The MB is one logical function of the MPN. It consists of media-processing equipment that terminates voice communications from a terminal in the form of Real-time Transport Protocol (RTP)[*15] and provides SCN with a variety of value-added media functions. It provides an API that abstracts supplementary media processing such as call-recording control and conferencing control for a service scenario. This abstract API is called the Media Composition Enabler (MCE).

Here, http/Simple Object Access Protocol (SOAP) is used for achieving communication between SCN-AS and MCE. SOAP is a protocol primarily designed for linking distributed services. Application developers find it advantageous because of its extendibility and ease of implementation.

When providing a service via the SEN infrastructure, SCN uses the receiving of a service-launch signal from SCIM as a trigger for launching a service scenario through the enabler section. The service scenario calls conditional-branching logic and various abstract APIs. This mechanism makes possible call-processing services such as a call-recording service and over-the-phone translation service.

4) ECN

The ECN consists of interface equipment for accommodating enablers. From the viewpoint of a service scenario that uses enablers, ECN provides a function for accessing those enablers without having to worry about

*12 **Java® EE 5**: The name and specification for a server-side Java development and run-time environment. In earlier versions it was called J2EE 1.2, but since version 1.5, the name has been changed to JavaEE 5. Oracle and Java are registered trademarks of Oracle Corporation, its subsidiaries, and affiliates in the United States and other countries. Company and product names appearing in the text are trademarks or registered trademarks of each company.

*13 **Interworking**: Interaction between communications systems.

*14 **Diameter**: An extended protocol based on Remote Authentication Dial In User Service (RADIUS), and used for authentication authorization and accounting in IMS.

the installation protocol or processing ability of each protocol. In the past, conventional telecom systems featured high interdependency between applications and a mixture of many protocol technologies and data formats. This made it necessary to support different protocols on communication nodes and different interfaces corresponding to node characteristics. The ECN was introduced to solve this issue. It supports multiple protocols (Representational State Transfer (REST), SOAP, HTTP) and includes node selection logic and a function for managing connection states. As such, ECN makes it easy for service scenarios to use modularized enablers in the provision of services.

2.3 Overall Configuration

Connection routes before and after introducing the SEN infrastructure are shown in **Figure 5**.

When applying supplementary services using the SEN infrastructure to voice calls, the IMS infrastructure connects with SCIM/SCN on the SEN infrastructure when the user makes or receives a call. This format conforms to the IMS Service Control (ISC) interface that connects to AS from S-CSCF. It may be possible to use the same type of connection route even if a new access network is introduced in the future. Access from IMS to SEN can be performed from both the originating network and terminating network, which means that either supplementary services launched by the originating user or supplementary services launched by the terminating user can be provided. The use of enablers is initially being restricted to service scenarios within the network. There are plans, however, to generalize the format for accessing enablers by introducing a Gateway (GW)[*16] to support future versions of the Internet and enable the use of enablers from 3rd party[*17] suppliers.

3. Examples of Application Services

3.1 Call-recording Service

The call-recording service is oriented to corporate users. It enables a subscriber to record the contents of a call when originating or terminating a call and to forward the file containing the recording to the subscribing company. A sales manager or sales representative, for example, may wish to record calls made by mobile terminal while away from the office so as to leave a trail of those business dealings. The call-recording service meets such a need—it records the contents of an originating or terminating call for the contract phone number in question and saves the call on server facilities belonging to the corporate subscriber. The conventional method for recording a voice call is to use a recording system installed within the corporate user's facilities. The aim of providing a recording system within a communications operator's system is

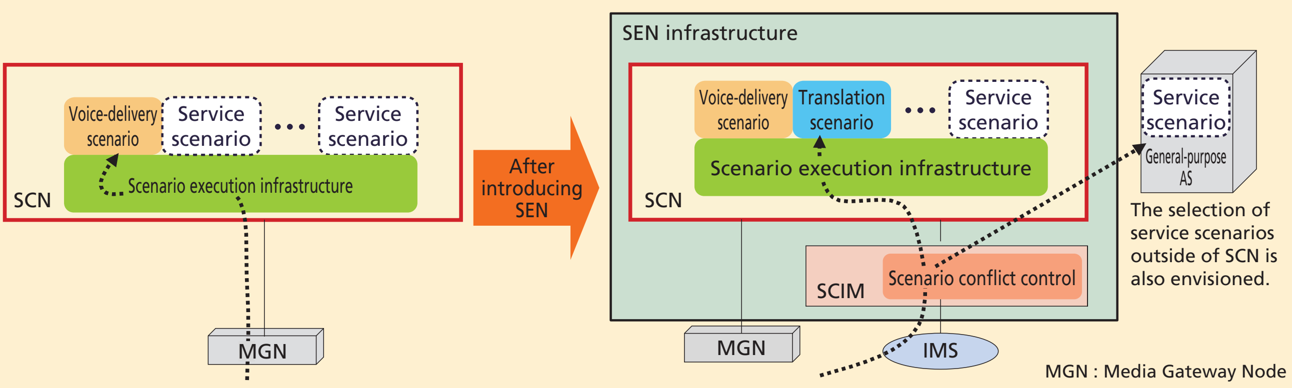

Figure 5 Connection routes before and after introducing SEN

*15 **RTP**: A protocol defined by IETF for real-time distribution of audio, video or other such media.

*16 **GW**: An intermediate device that has functions such as protocol conversion and data transfer to allow communication between devices.

*17 **3rd party**: A maker that is not in partnership with the provider of a certain product and that creates software for that product from disclosed specifications.

to enable user facilities related to recording to be downsized and to provide a service model featuring a small up-front investment.

1) Service Implementation Functions

- Subscribing/charging functions

  To enable only subscribers to the call-recording service to use the service, a call-recording contract is concluded as a supplementary contract to the telephone-number contract. Charges are applied to that telephone number based on the call-recording contract.

- Call-recording function

  This function records and temporarily stores on NTT DOCOMO facilities a voice call either originating from or terminating with the call-recording subscriber.

- Recorded-data transfer function

  This function transfers recorded data whenever obtained to the server facilities of the corporate subscriber associated with the telephone number in question. Recorded data is deleted after being successfully transferred.

- Corporate profile submission function

  This function enables an Account Manager (AM) of a subscribing company to set a profile for each telephone number, corporate subscriber and transfer group. The AM provides information on the corporate facilities to which recorded data are to be transferred and sets active/inactive states for recording. The AM can also check the file-transfer state for recorded data and can force a file-transfer operation in the event of a file-transfer error (file-transfer state management function).

2) Service Implementation Method

The architecture and functions of the call-recording service are shown in **Figure 6**.

- Call-recording function

  The SEN infrastructure is used here to perform voice-call processing. In IMS, a voice call is usually achieved by a direct connection between terminals, but in the case of call recording, the processing is performed at the MB, which means that the voice call must be drawn into the MB. To be more specific, as SCIM is registered in the SiFC of S-CSCF, an Invite[*18] signal will necessarily be sent to SCIM on the occasion of a call originating from or terminating with the call-recording subscriber (Fig. 6 (1)). The SCIM now determines the existence of a call-recording contract based on contract information and checks for any service conflicts (Fig. 6 (2)), and launches a service scenario (call-recording scenario) stored on SCN. Next, based on an Invite signal from SCIM, SCN creates a conference room within MB, launches a conference service with the originating and terminating parties as participants, and draws the voice call into MB. When the call begins, MCE within MB is instructed to output guidance on initiating the recording and to then record the call (Fig. 6 (3)). In short, MCE outputs guidance and records voice data based on the service-scenario request.

- Recorded-data transfer function

  The file-transfer enabler transfers recorded data as instructed by the call-recording service scenario. This process allows that data to be encrypted (Fig. 6 (4)).

- Corporate profile submission function

  Corporate profile data is stored in the SCN database (Fig. 6 (5)). Here, to enable an AM to submit profile data, a Web interface is provided on a specific terminal (client) that can access SCN from a corporate LAN. In this way, an AM can perform a variety of profile operations and review profile data using a browser-based GUI[*19].

3.2 Over-the-phone Translation Service

NTT DOCOMO's over-the-phone translation service translates a voice call in Japanese into a foreign language (English, Chinese or Korean) and vice versa. It can incorporate the results of translation in the voice call to achieve communication between remotely

*18 **Invite**: An SIP signal requesting a connection.

*19 **GUI**: A user interface that uses many icons and other graphics on which basic operations are performed mainly with a pointing device.

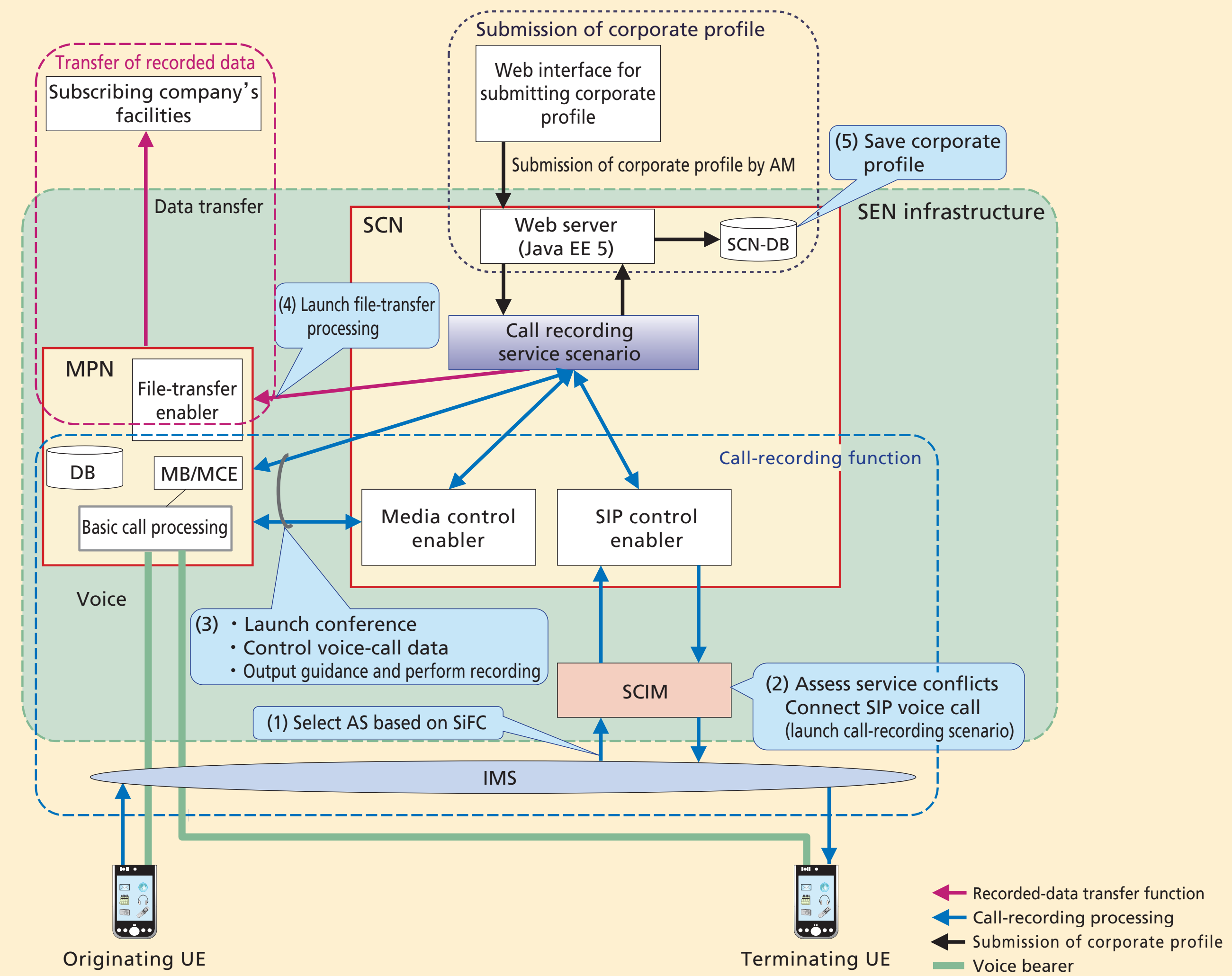

Figure 6 Architecture of call-recording service

located people speaking different languages (remote interpretation function). An auxiliary tool for controlling outgoing transmissions is provided in the form of an Android™*20 application that displays the results of voice recognition and translation on the user's screen. Additionally, to support non-remote, face-to-face communication between people speaking different languages, voice data input by microphone on one terminal can be translated using an over-the-phone translation application (face-to-face interpretation function).

The over-the-phone translation service has been provided on a trial basis since November 2011 using 3rd Party Call Control*21 and a center-originating/user-terminating call system. In this development, however, it will be possible for the user to connect with this service by simply prefixing the phone number of the other party with a special number. This will enable users to make use of this service by operations that are similar to those used for ordinary calls.

1) Service Implementation Functions

- Remote interpretation function

The over-the-phone translation service is achieved by five key functions. These are a function for drawing the voice call into the operation center, a function for recog-

*20 **Android™**: An open source platform targeted mainly at mobile terminals or promoted by Google Inc., in the United States. Android™ is a trademark or registered trademark of Google Inc., in the United States.

*21 **3rd Party Call Control**: A technology that enables a 3rd party to control the connections among two or more terminals.

nizing a user instruction to commence translation, a function for performing voice recognition against the above voice call, a function for translating the results of voice recognition, and a function for subjecting the results of translation to speech synthesis and incorporating the synthesized speech in the voice call. For users having the Android application, over-the-phone translation achieves CS/PS linking[*22], which in an example of "creating new value through telecom/Web convergence," a key concept of SEN. Providing a function for displaying the results of voice recognition and translation makes for a service that is even easier to use.

- Face-to-face interpretation function

This function adopts an over-the-phone translation application on one terminal to subject voice data picked up by a microphone to voice-recognition processing and return translation results.

2) Service Implementation Method

The architecture and functions of the over-the-phone translation service are shown in **Figure 7**.

- Remote interpretation function (network recording)

The user launches the over-the-phone translation service by prefixing the other party's phone number with the special number 138 and a two-digit language code and then making the call (Fig. 7 (1)). The number 138 indicates launching of the over-the-phone translation service and the subsequent two-digit code signifies the other party's language to be interpreted. However, we note here that users operating

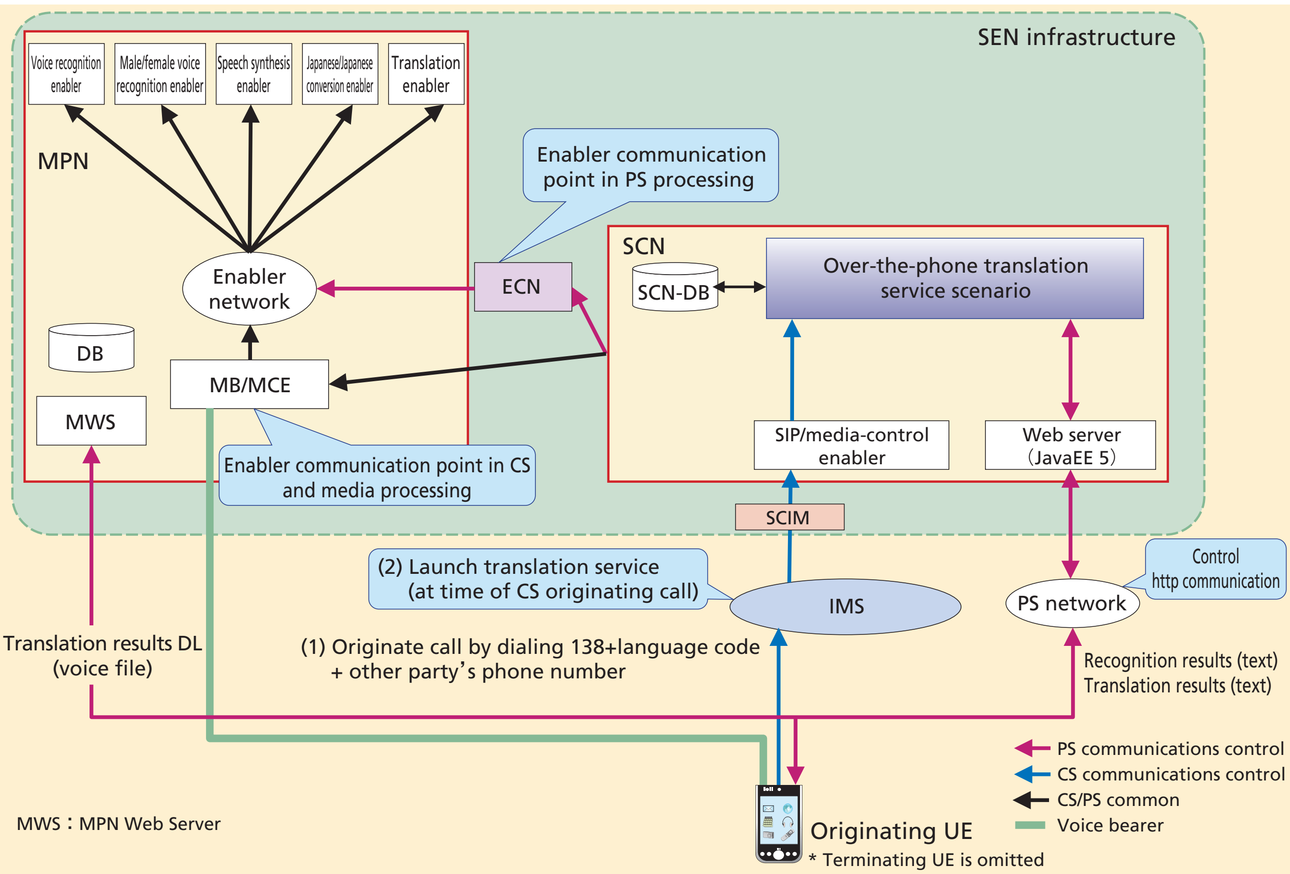

Figure 7 Architecture of over-the-phone translation service

*22 **CS/PS linking**: The linking of CS voice communications and PS data communications.

the Android application on User Equipment (UE) can originate a call using this service without having to worry about the special number or language code.

An existing AS within IMS recognizes this special number (138) as a SEN-service number that requires a connection with SCIM to be made. After checking for service conflicts, SCIM connects the call with the appropriate SCN service scenario (over-the-phone translation scenario) (Fig. 7 (2)). The method used here for drawing the voice call into media processing equipment is the same as that of the call-recording service.

The various components of the enablers required to interpret speech during a voice call are called "engines" as shown in **Table 1**. From the viewpoint of a service scenario, all of these engines are consolidated in MCE, which actually absorbs the interface differences among these engines. This scheme makes it possible to specify abstract media processing instructions from a service scenario. It also enables engines to be upgraded as needed without affecting service scenarios.

- Face-to-face interpretation function (terminal recording)

The face-to-face interpretation function provides a service that uses only packet communications (no voice call). The entire service is controlled using http signals between the over-the-phone translation application and the SCN service scenario. Based on instructions received from the scenario, the terminal directly connects to a media enabler (voice recognition) via ECN on the SEN infrastructure without having to use the scenario as an intermediary. The same enabler can be used from another server by opening up the API for that enabler via ECN.

- Application launch function

The SCN over-the-phone translation scenario receiving a call connection for over-the-phone translation sends an SMS message to originating/terminating users instructing the application to start up. The over-the-phone translation application receiving this message then starts up enabling the service to be used.

Table 1 List of engines

Name	Description
Voice recognition engine	Inputs voice data and converts utterances to text
Male/female voice recognition engine	Inputs voice data and determines whether voice is male or female; this result determines what gender to use in speech synthesis (the result of interpreting a male is a male voice and the result of interpreting a female is a female voice)
Speech synthesis engine	Generates speech using the results of translation as input; the speech so generated is male or female depending on the results of male/female voice recognition
Japanese/Japanese conversion engine	Analyzes Japanese input for proper nouns as preprocessing to the translation process; identifying and removing proper nouns enables a more accurate translation to be performed
Translation engine	Translates input to another language

4. Conclusion

This article described an infrastructure for achieving a service enabler network to facilitate the quick, cost-efficient development of value-added services on the network. The architecture described here enables multiple enablers to be combined so that services can be provided in a flexible manner by mostly developing service scenarios. Examples of applying the SEN infrastructure to provide a service were also presented.

Looking forward, we plan to introduce gateways to the outside with the aim of providing enabler functions as APIs to the Internet and external, 3rd-party networks.

REFERENCES

[1] K. Inoue et al.: "SCN: Providing Service Scenario Execution Infrastructure Using Abstract APIs—Koe-no-Takuhaibin—," NTT DOCOMO Technical Journal, Vol. 13, No. 2, pp. 20-28, Sep. 2011.

[2] A. Miyata et al.: "Media Processing Node for Providing Value-added Media Services," NTT DOCOMO Technical Journal, Vol. 11, No. 1, pp. 4-12, Jun. 2009.