# High Audio Quality Melody Call for VoLTE

**Core Network Development Department**
**Mana Kaneko, Yu Kojo**†
**Corporate Strategy & Planning Department**
**Hiroaki Nakajima**
**Online Marketplace Department**
**Susumu Ogata**

## 1. Introduction

In June 2014, NTT DOCOMO began the first voice over LTE (VoLTE) service in Japan, enabling users to call each other with high quality audio. Because users must have a VoLTE-capable terminal to enjoy high audio quality communications with each other, the number of users who have VoLTE-capable terminals and opportunities for users to enjoy high audio quality calling need to be increased. Thus, to promote VoLTE terminal usage, we have developed a high audio quality DOCOMO Melody Call®[*1] service, which became available in October of 2014. This service enables VoLTE terminal users to experience high audio quality regardless of the type of terminal the other party is using. Melody Call is a service that plays music or voice selected by the receiver for the caller, in place of a Ring Back Tone (RBT)[*2]. NTT DOCOMO began providing this service in September of 2003.

This article discusses the background to the development of high audio quality Melody Call, an overview of the service, and describes how this service was achieved and the steps taken during development.

## 2. Development Background

For the codec[*3], in place of the Adaptive Multi-Rate NarrowBand codec (AMR-NB)[*4] [1], VoLTE enables high audio quality calling using the wide band Adaptive Multi-Rate WideBand codec (AMR-WB)[*5] [2]. There are two objectives for raising the quality of Melody Call audio. The first is to make VoLTE more appealing, while the second is to raise the added value of the Melody Call service for future service developments.

• Appealing to VoLTE

The high audio quality of VoLTE is only available when both the caller and receiver are using VoLTE terminals. For this reason, even if users purchase a VoLTE terminal, the opportunities to experience high audio quality calling will be limited to calling other users who are also using VoLTE terminals. To counter this situation and increase the opportunities for users to experience high audio quality even if the person they are calling does not have a VoLTE terminal, we developed the high audio quality Melody Call as a mechanism for raising awareness of the benefits of the high audio quality of VoLTE.

† Currently Solution Business Department

*1 **Melody Call®:** An NTT DOCOMO service that lets users replace the mobile phone ring back tone with their favorite music. It is a registered trademark of NTT DOCOMO.
*2 **RBT:** A ring back tone is played via the network to the caller while the receiver is being called (ring ring, ring ring etc).

- Increasing the added value of Melody Call

Melody Call [3] has been available since the feature phone era, and is a service currently used by many users. By developing the Melody Call service for high audio quality available with VoLTE, this development aims to raise the appeal of VoLTE with improved audio quality (music audio quality), and thus further promote usage of the service among existing users while acquiring new users.

## 3. Service Overview

High audio quality Melody Call is available under the following conditions:

- Usage terminals and usage areas

To enjoy high audio quality Melody Call, callers must use a VoLTE terminal and be within an LTE coverage area. Call receivers must be subscribed to Melody Call, and have high audio quality music set for it (**Table 1**, **Figure 1**).

- Rates

High audio quality does not incur any changes to rates - existing users can use high audio quality Melody Call at their current rates.

---

*3 **Codec:** Technology for coding and decoding data such as audio signals.
*4 **AMR-NB:** A 3GPP standard conventional audio encoding method.
*5 **AMR-WB:** A 3GPP standard wideband audio encoding method.

**Table 1 Conditions for providing high audio quality Melody Call**

| Caller | Receiver | Melody Call subscription | Music settings | Ring back tone | (Reference) Call |
|---|---|---|---|---|---|
| VoLTE terminal user (in LTE area) | VoLTE terminal user (in LTE area) | Yes | High audio quality music | High audio quality Melody Call | High audio quality calling |
| | | | Existing music settings | Conventional audio quality Melody Call | |
| | | None | — | RBT (ring ring etc) | |
| | VoLTE terminal user (in 3G area), or non-VoLTE terminal user (in 3G/LTE area) | Yes | High audio quality music | High audio quality Melody Call | Conventional audio quality calling |
| | | | Existing music settings | Conventional audio quality Melody Call | |
| | | None | — | RBT (ring ring etc) | |



- A terminal that supports high audio quality codecs (a VoLTE terminal) must be used

DOCOMO Network

- Must have a Melody Call subscription
- Must have high audio quality music set.

LTE area
Caller

- Must be in an LTE coverage area
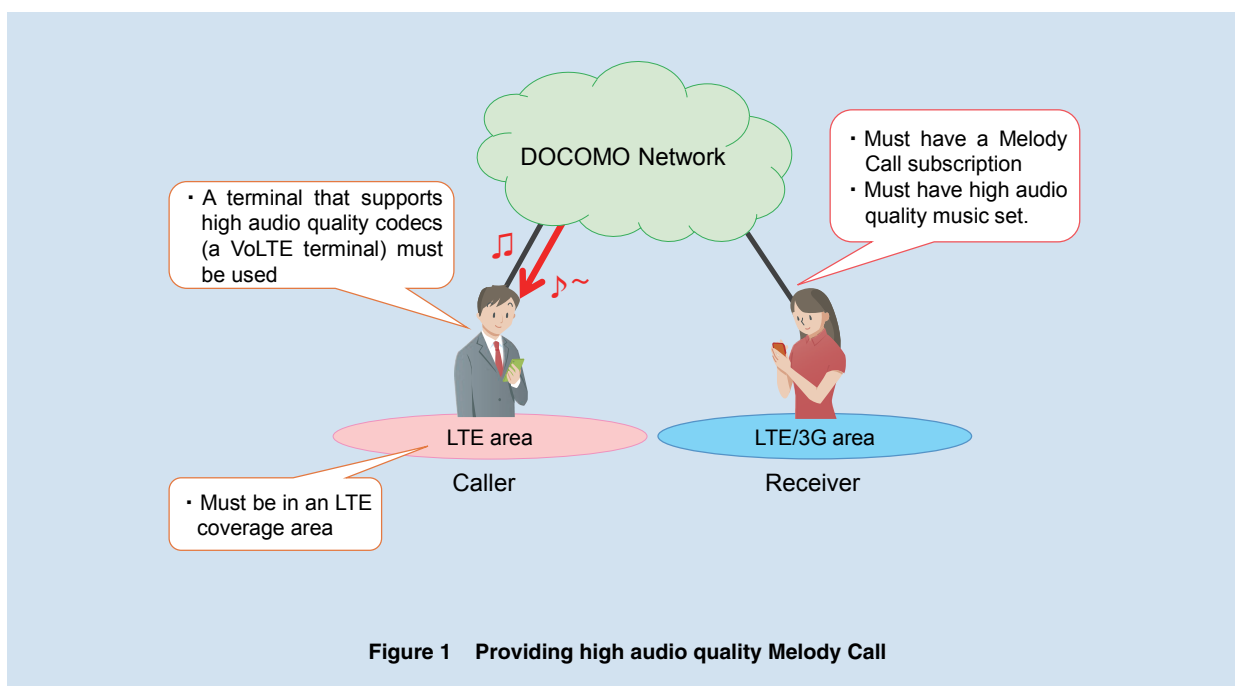
LTE/3G area
Receiver

**Figure 1 Providing high audio quality Melody Call**

• Music settings

   Future plans include high audio quality music to be offered by Content Providers*6 for Melody Call as standard, and a gradual switch to automatic provision of high audio quality sources of the music currently set by users.

• Trial listening

   Trial listening to Melody Call sounds will also be available in high audio quality. This will enable VoLTE terminal users to listen to samples of high audio quality music on Content Provider sites or the Melody Call setting site.

• Trial dialing

   VoLTE terminal users in LTE areas will be able to try high audio quality Melody Call by dialing the 157005 VoLTE trial number, or high audio quality calling by dialing 157001 (As of January 2015).

## 4. Implementing Method

### 4.1 Basic Call Processing
### (High Audio Quality Codec Selection Logic)

   The following describes the network architecture and process flow for high audio quality Melody Call (**Figure 2**).

(1) The network receives a call request (SIP_INVITE) from the caller, and then delivers the call request (SIP_INVITE) to the receiving terminal.

(2) The terminal that received the call request pages the receiver, and sends a paging complete response (SIP_180 Ringing) to the network.

(3) The Application Servers Node (ASN) that receives the paging complete response (SIP_180 Ringing) judges whether the receiver is subscribed to Melody Call from receiver subscription data, and then sends a Melody Call output request (SIP_INVITE) to the Media Processing Node (MPN)*7 that is operating as the Melody Call Application Server (AS)*8.

(4) The MPN that receives the Melody Call output request (SIP_INVITE) judges whether the calling terminal is requesting a high audio quality call from the SDP*9 data[5] included in the signal.

(5) The MPN acquires the sound source set for the receiver from storage.

(6) The MPN encodes the acquired sound source with

---

*6 **Content Provider:** A business that offers Melody Call sound sources.
*7 **MPN:** A node of the NTT DOCOMO core network. It currently provides various media services including voice answering, Melody Call and other voice media services; video media services such as videophone answering; and SMS.
*8 **AS:** A server that executes an application to provide a service.



**Figure 2　NW architecture and process flow**
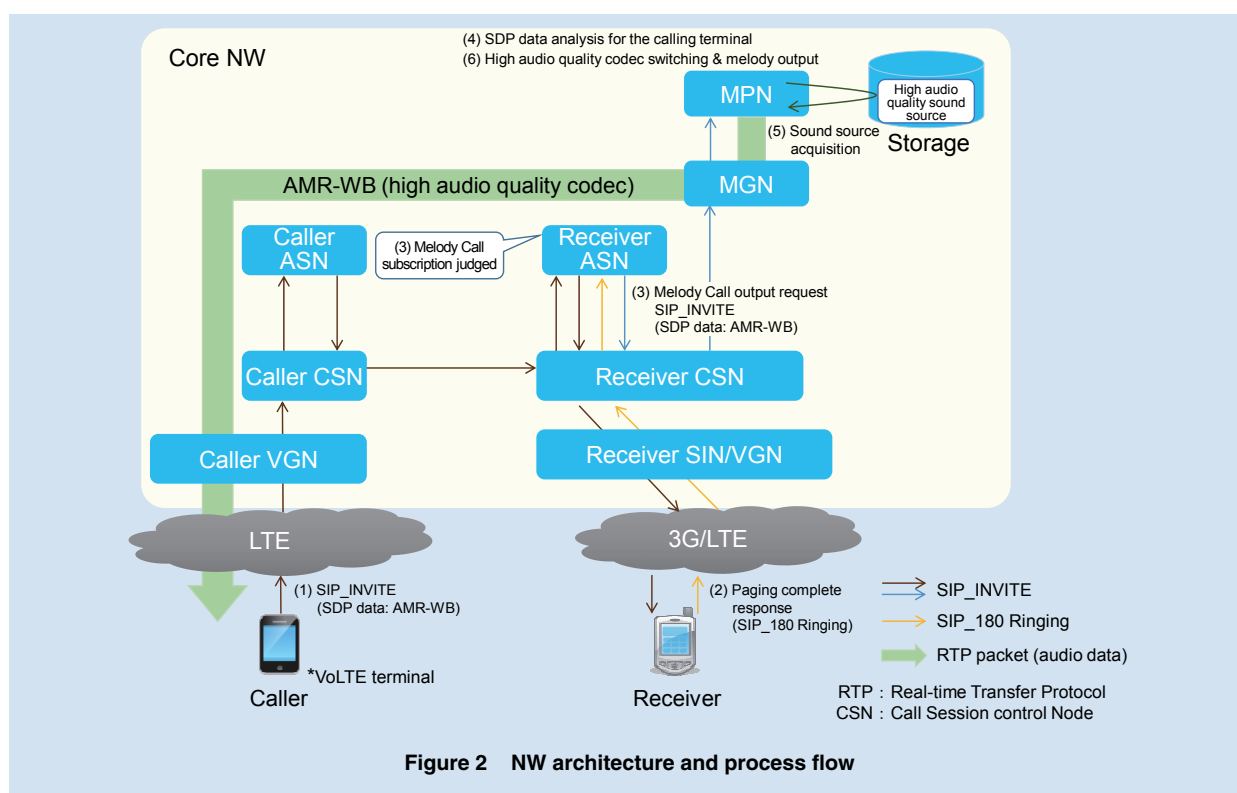
a high audio quality codec (AMR-WB), and outputs it to the calling terminal so that the high audio quality Melody Call is replayed on the calling terminal. Acquiring sound source from storage and switching to high audio quality codec is performed upon reception. Also, if the calling terminal does not support AMR-WB, codec conversion is performed to match the calling terminal.

This processing is basically the same for currently offered Melody Call[3]. The main difference with existing processing is step (4), judging SDP data in the caller signal, and step (6) outputting Melody Call with the high audio quality codec.

## 4.2  LTE to 3G Handover Actions

When moving into a 3G area during VoLTE voice calling, calls are continued using Single Radio Voice Call Continuity (SRVCC) technology. NTT DOCOMO has been providing SRVCC since implementing VoLTE [8] so that the caller can continue to hear the sound source even when moving into a 3G area while listening to the RBT or Melody Call. However, with handover to 3G through SRVCC, the codec for the music being played must be switched from the high audio quality Melody Call codec to the conventional audio quality AMR-NB codec.

The following describes SRVCC processing during Melody Call output (**Figure 3**).

(1) When SRVCC is executed, the VoLTE Gateway Node (VGN) sends notification (SIP_UPDATE) to switch from LTE to 3G to the Call session Control Node (CSN). CSN transfers SIP_UPDATE to ASN.

(2) The ASN that receives SIP_UPDATE sends notification (SIP_ReINVITE) to switch from LTE to 3G to the Media Gateway Node (MGN).

(3) The MGN that received SIP_ReINVITE analyses ReINVITE, and converts the AMR-WB received from MPN to AMR-NB and sends it.

## 4.3  Adoption of High Audio Quality Sound Sources with an Eye to Next-generation Audio Codecs

Sound sources currently in storage are compressed with a quality lower than AMR-WB, therefore, sound sources must at least be upgraded to AMR-WB or equivalent to enable callers to hear high audio quality Melody Call. Sound source upgrade is also described in Chapter 3, and there are plans for content providers to gradually replace sound sources with high audio quality sound sources in future (**Figure 4**).

*9  **SDP:** A protocol to describe information such as IP addresses necessary for initiating sessions in the IMS. It is also used to describe session information relating to SIP, a call control protocol.
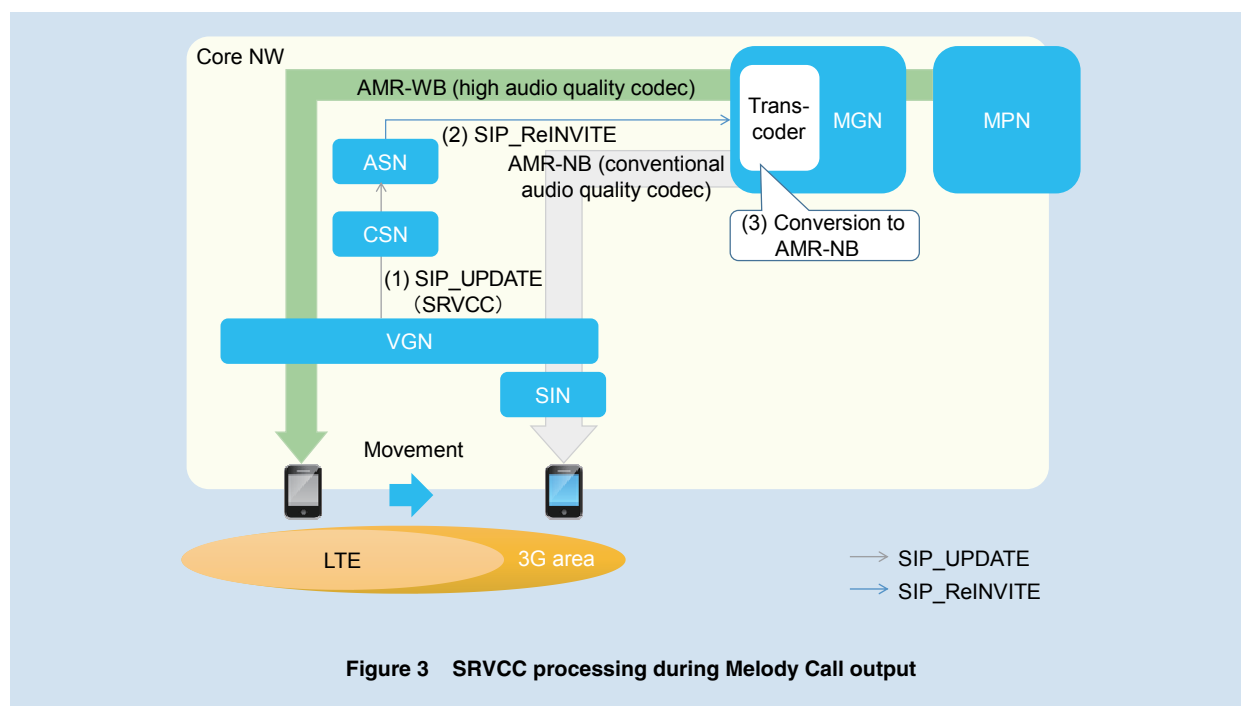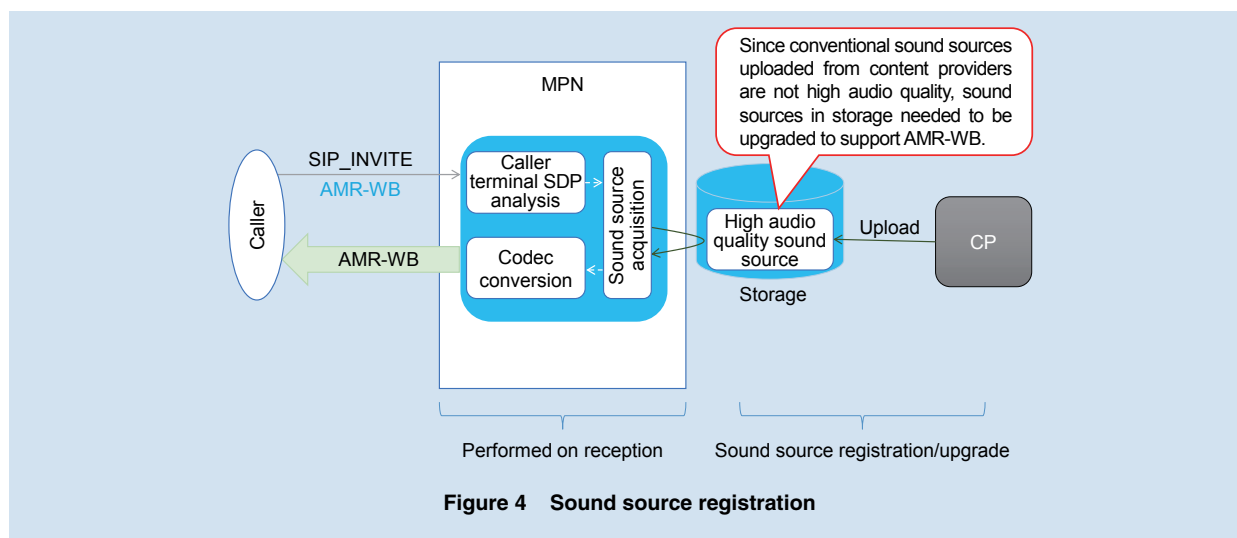


**Figure 3  SRVCC processing during Melody Call output**

**Figure 4  Sound source registration**

Currently, Enhanced Voice Services (EVS)*10 [6] [7] is gaining attention as the next-generation audio codec, thus, with an eye to further increases in audio codec performance and in consideration of bit rate and audio range, high audio quality sound sources compatible with EVS have been adopted. These high audio quality sound sources will be compatible with next-generation Melody Call codecs and will not require upgrade (**Table 2**).

## 5.  Conclusion

This article has described the high audio quality Melody Call development accompanying the implementation of VoLTE, and an overview of the service and methods of delivery. Into the future, we aim to continue to make DOCOMO services even more convenient.

REFERENCES

[1]  3GPP TS26.071: "Mandatory speech CODEC speech processing functions; AMR speech Codec; General description."
[2]  3GPP TS26.190: "Speech Codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; Transcoding functions."
[3]  H. Hayasaka et. al: "Development of Melody Call Service," NTT DOCOMO Technical Journal, Vol. 12, No. 1, pp. 64-69,

**Table 2  Sound source quality and codec performance**

| Audio codec | Conventional audio quality Melody Call sound source | High audio quality Melody Call sound source |
|---|---|---|
| **AMR-NB** (Bandwidth: 300 Hz to 3.4 kHz/bit rate: 4.75 to 12.2 kbps) | Y | Y (Conventional audio quality) |
| **AMR-WB** (Bandwidth: 50 Hz to 7 kHz/bit rate: 6.6 to 23.85 kbps) | — | Y |
| **EVS** (Bandwidth: 20 Hz to 16 kHz/bit rate: 5.9 to 128 kbps) | — | Y |

Apr. 2004 (Japanese only).
[4]  Y. Shimada et. al: "IP-based FOMA Voice Network toward Enhanced Services and Improved Efficiencies," NTT DOCOMO Technical Journal, Vol. 12, No. 1, pp. 4-14, Jun. 2010.
[5]  3GPP TS24.229: "IP multimedia call control protocol based on Session Initiation Protocol (SIP) and Session Description Protocol (SDP); stage3."
[6]  3GPP TS26.114: "IP Multimedia Subsystem (IMS); Multimedia telephony; Media handling and interaction."
[7]  K. Tsutsumi et. al: "3GPP EVS Codec for Unrivaled Speech Quality and Future Audio Communication over VoLTE," NTT DOCOMO Technical Journal, Vol. 16, No. 4, pp. 6-13, Apr. 2015.
[8]  K. Tokunaga et. al: "VoLTE for Enhancing Voice Services," NTT DOCOMO Technical Journal, Vol. 16, No. 2, pp. 4-22, Oct. 2014.

***10  EVS:** A 3GPP standardized next-generation audio encoding method.