FORSCHUNG

5G MIXED-MODE MULTICAST/ BROADCAST RELEASE 17

A NEW APPROACH TO SCALABLE MEDIA DISTRIBUTION USING 5G

BY JORDI J. GIMENEZ (IRT)

Over-the-Top (OTT)-Plattformen wie die Mediathek und die Audiothek bieten lineare und On-Demand-Dienste für mobile Nutzer an. Die mangelnde Skalierbarkeit von Unicast-Verbindungen macht es allerdings schwierig, einen nachhaltige Quality of Service zu gewährleisten, insbesondere bei Events mit großem Publikum. 5G bietet mit der Integration von Multicast- und Broadcast-Fähigkeiten die Option, mobile Netzwerke zu entlasten und eine effizientere Nutzung von Funkressourcen zu ermöglichen.

Over-the-top (OTT) platforms such as the Mediathek and Audiothek are currently carrying linear and on-demand services to serve mobile users. The lack of scalability of unicast connections makes it difficult to provide a sustained Quality of Service, especially for large audience events. 5G is starting to integrate multicast and broadcast capabilities with an opportunity to offload mobile networks and provide a more efficient use of radio resources.

The delivery of TV and radio services to smartphones, tablets and, in general, connected devices is currently done with over-the-top (OTT) platforms such as the Mediathek and Audiothek. These platforms suffer, however, from the lack of scalability as the content is delivered using unicast connections. This is obvious for on-demand content as every user can make a selection of the desired programme at the specific time they want to consume it. Multiple independent connections are unavoidable although some efficiency might be possible by means of caching mechanisms.

However, when it comes to live or linear content, where users connect and receive exactly the same programme at the same time, possibilities for improving delivery efficiency are worth to be considered. The use of multicast is an option being already explored in the context of fixed networks. With HTTP Adaptive Bitrate (ABR) currently struggling to provide scalability and stable quality of experience comparable to that in traditional broadcast platforms, Multicast ABR and ultra-low latency applied to IPTV delivery offer a unique opportunity to improve user's quality of experience. Mobile networks might need to evolve in a similar direction to allow media delivery to scale according to e.g. the number of concurrent users in a cell, while guaranteeing a desired and sustainable quality of service. The situation becomes even more challenging than for fixed networks given the scarcity of radio resources, which need to be shared by many users, and additional issues such as interferences from other cells or the need to guarantee mobility among them. Although the problem might not be there yet as TV consumption over mobile networks is onerous for consumers' pockets, the changing environment with more unlimited data contracts or zero-rating offers may create an unavoidable scenario. Video streaming habits are changing with OTT platforms not only being used for on-demand consumption but also becoming reference platforms for the consumption of live events such as sports and especially among young audiences [1].

3GPP is working on solutions under the umbrella of 5G with two main tracks to address media distribution: Terrestrial Broadcast, with the well-known 5G Broadcast standard which deserved an article in the previous number of FKT [2], and the 5G Mixed-Mode Multicast/Broadcast, a new alternative that is just starting to be developed and that is explained here.

Media Distribution in 3GPP

LTE and 5G networks are and will be able to support the distribution of media services in the form of the current offers provided by applications such the Mediathek or Audiothek, with a combination of linear (e.g. current TV and radio services) and non-linear (e.g. on-demand, podcasts) services. However, the way these platforms reach the final user and how users perceive and enjoy the content originally prepared by broadcasters is unpredictable, which suppose a big gap with respect to traditional broadcast technologies where an end-to-end control of the quality of service can be guaranteed.

Mobile technologies can be shaped in order to provide a managed distribution of media services, which goes beyond the current over-the-top (OTT) models, see Fig.1, and enables third-party service providers and mobile network operators to establish several arrangements. 3GPP defines in TS 22.101 [3] the support of Television (TV) services in 3GPP so that networks are able to provide unicast and broadcast transport to support the distribution of TV programs. This considers TV services as a whole (free-to-air and subscribed services) for which different requirements and, therefore, solutions are specified. The document also expresses interest from the TV industry to be able to deliver TV services to connected consumer devices (smartphones, tablets and TVs) anywhere, by using cellular networks, from over-the-top (OTT) service

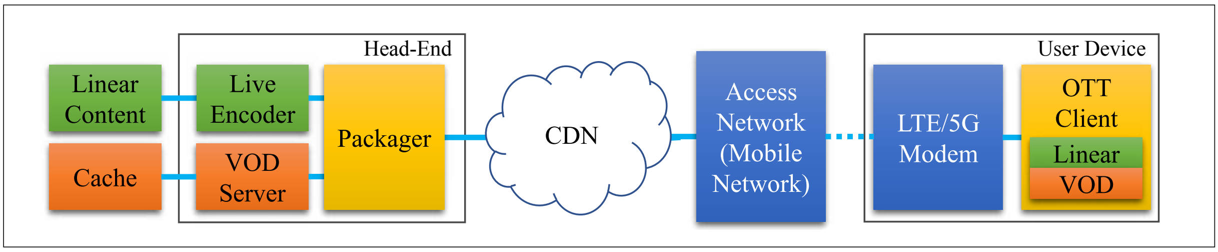

Figure 1. Over-the-top (OTT) media distribution to mobile devices

providers up to enabling free-to-air TV. From that point, 3GPP has been developing different solutions. In particular, linear TV services can be delivered already with LTE using eMBMS (evolved Multimedia Broadcast Multicast Service) or Packet Switched Streaming (PSS) architectures [4].

The PSS architecture, originally developed for 3G and 4G with the aim of providing a solution for MNOs operating their Mobile TV offers, is a closed solution where services are under control of the operator. Despite being evolved for a better integration with new codecs and protocols such as DASH (Dynamic and Adaptive Streaming over HTTP) and HEVC (High Efficiency Video Coding), the PSS architecture was found not to be optimal for its integration with current online video services.

On the other hand, eMBMS also introduces key concepts to enable the distribution of services through mobile networks where the service provider holds the control on many parts of the system. It permits to offload the network when content is concurrently consumed by many users. These features are already well-known as eMBMS is the underlaying system of what is popularly known as 5G Broadcast [5]. Beyond the modifications required for the provisioning of free-to-air broadcast services, in the Rel-14 version of the 3GPP specifications, eMBMS also includes an interface from the content provider to the MBMS system (xMB interface) as well as an MBMS-API between the client and the application at the user device. In addition, the shared broadcast capability enables the distribution of broadcast services through MNO networks with a common entry point for the service provider and an infrastructure sharing approach between different MNOs.

Leveraging the high potential of 5G enhanced mobile broadband connectivity with increased data rates and low latency, and new network features such as network slicing or edge computing, 3GPP is developing a media architecture fully integrated within the 5G system. This aims at supporting the most recent advancements in terms of media and video content providing augmented quality of service for traditional audio and video services as well as emerging formats for virtual/augmented/mixed reality. The 5G Media Streaming Architecture (5G-MSA) [6] is the nuclear system to enable different business arrangements between on-line media service providers, broadcasters, and mobile network operators. With the 5G-MSA, network and device functionalities are exposed to third-party providers enabling the use of 5G capabilities in the best way to ensure an increased quality of service for connected users. The new architecture is a reality in 3GPP from 5G Release-16 and is currently supporting unicast downlink media distribution as well as uplink streaming. The later, an interesting option for media production and contribution scenarios.

The 5G-MSA introduces the concept of trusted media functions, see Fig. 2, which are implemented in both the network and the user device and also define APIs to interface with external media servers and functions. This effectively means that functions commonly deployed outside the network domain can be integrated within it. ABR encoders, streaming manifest generators, segment packagers, Content Delivery Network (CDN) servers and caches, Digital Rights Management (DRM) servers, content servers for advertisement replacement, manifest modifications servers, or even metrics servers can now be allocated within the 5G network to improve the delivery of the service. For instance metrics collection and reporting may provide information related to the user experience, streaming sessions can be monitored on the user device and reported back to the service provider or even to the network where the information may be used potential transport optimization within the mobile network.

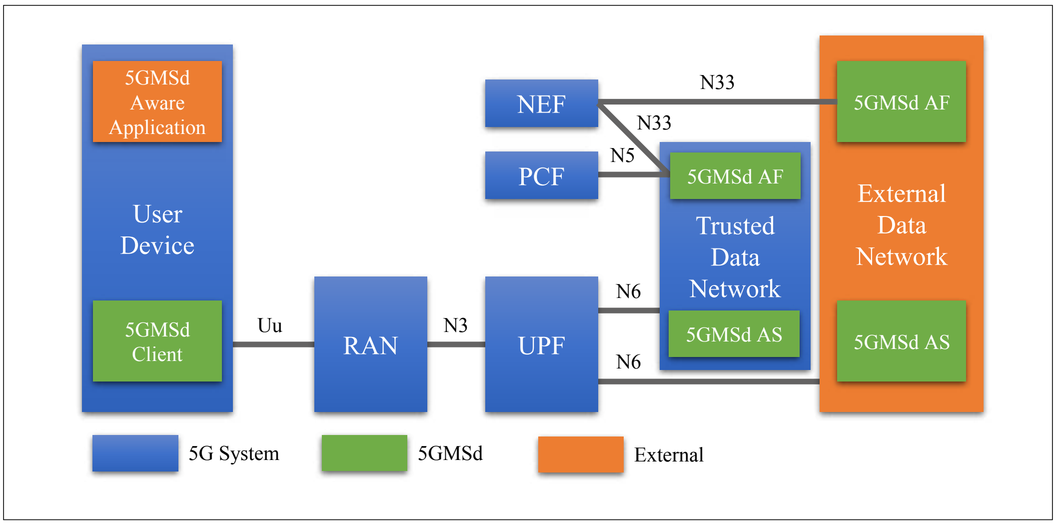

Figure 2. 5G Media Downlink Streaming (5GMSd) for the 5G System [7]. NEF (Network Exposure Function), PCF (Policy Control Function), UPF (User Plane Function) are network functions of the 5G System. AF (Application Function) and AS (Application Server) are defined by the 5GMSd.

This opens the door to use 5G not just as a better 4G network but as an option shaped to cover needs and resolve problems that IP distribution via mobile networks is currently facing. Of course, without preventing the traditional delivery using an OTT approach.

A role for multicast/broadcast also in 5G

The lack of scalability is one of the issues that derives in a poor quality of service while consuming live OTT services. There are different segments in the delivery path that may affect the system to perform in the expected way, not always related to the access network (the last mile). Potential dimensioning and processing problems at the origin server as well as the CDN operator need to be resolved, but indeed various problems might arise in the mobile network operator domain. Investing in coverage and data traffic while providing enough performance and attractive subscription fees is challenging. Dimensioning the access network for the peak of consumption, which sometimes is hard to predict, is not an economic solution for MNOs. And here is where multicast and broadcast capabilities may play an important role.

Multicast and broadcast capabilities are still missing features of the new 5G system as the only possibility to use broadcast is based on LTE with elements not compatible with new architectures developed specifically under 5G. This means the work is still to be developed and, although this could be seen as a disadvantage, it can bring the possibility to develop a new approach to multicast and broadcast from scratch. Certainly, without proper mechanisms that can make the system scale according to demand, any possibility to provide sustainable quality for connected users consuming massive audience programmes is doubtful.

3GPP originally developed MBMS for 3G enabling video broadcasting and streaming services and later introduced eMBMS for LTE. With LTE-based 5G Terrestrial Broadcast (5G Broadcast), eMBMS has been further developed to support TV and radio services in the way terrestrial broadcast standards (e.g. DVB) work. This work is based on LTE and 3GPP already recognizes in [8] that a simpler architecture developed from scratch for dedicated broadcast could provide benefits considering some of the pre-existing functional properties of 5G such as the SBA.

3GPP is considering the introduction of multicast and broadcast capabilities for the 5G system architecture in Release 17, initially targeting an architecture to fulfil requirements associated to IoT, Public Safety, V2X and dedicated broadcasting use cases (provision of linear TV, radio and hybrid services). At the radio access network (RAN), a new standardization activity is also started for Release 17 to specify the New Radio (NR) mixed-mode multicast/broadcast [9] which should initially involve the possibility to use multicast and broadcast at cell level or between a small group of cells i.e. not targeting those scenarios where 5G Broadcast has been designed to support (e.g. large area SFN areas, high-power high-tower deployments). The initial intention is to bring an option where the carrier conveys both unicast and multicast/broadcast content, in a mixed (mixed-mode) fashion that can benefit the improvements in the radio layer of 5G-NR, see Table I. It is also clearly stated that any of the current developments should prevent a future evolution of the system towards similar features as those in LTE-based 5G Broadcast.

How exactly the characteristics of the role of multicast and broadcast would be in 5G is still under discussion. Several models are confronted, from establishing an end-to-end multicast pipe where the network just acts as a transparent multicast transport network, to repeating the same model followed in eMBMS where it is possible to allocate functionalities within the core network that decide the forwarding of multicast/broadcast or unicast traffic to the end devices. In the first case, the application at the smartphone would need to communicate with the origin server in order to switch traffic from unicast to multicast/broadcast (i.e. all decisions are in the domain of the service provider), in the second case, there is a communication between the device and network functions that may decide how to deliver traffic to the user.

Whatever the solution is, one feature that has been recognized useful is the ability of dynamically switching users from unicast to multicast/broadcast and the way back. This is key to guarantee an efficient use of scarce radio resources, by making use of unicast, multicast or broadcast when they are really needed. This means in practice that users who start watching a live TV channel in the ARD Mediathek (e.g. Das Erste) will be connected using unicast in the way today this app works. When more users under the same cell or coverage area get connected to Das Erste this stream would be delivered using multicast or broadcast instead of unicast. Therefore, this avoids repeating the connection for every single user as a unique pipe will be established to serve many (a group) or all users, see Fig. 3. This has an enormous impact in the available radio resources. With unicast-only connections, a new independent unicast pipe is required per user what makes that, in the event of a highly popular programme (e.g. a football match), the network collapses and either more users are unable to connect to the TV channel or the quality of service of those which are already connected gets degraded, especially when adaptive bit-rate formats are used (i.e., less bit-rate per user is available). When the number of users concurrently consuming the same service reduces the users transit again from multicast or broadcast to unicast as serving them independently may not cause a big impact on the network.

This feature would look new but to some extent is already available in eMBMS (LTE) with a technology called MooD (MBMS Operation on Demand). Included as part of Release 12, MooD allows the network to count the number of users consuming the same unicast content and configure for eMBMS when this is detected and according to a threshold. Together with the service continuity mechanisms in eMBMS, it is also possible that users move across a mobile network

Table 1. Comparison between Capabilities of LTE and 5G-NR Radio Layer

	LTE	5G-NR
Frequency of Operation	Up to 6 GHz	Up to 52 GHz
Carrier Bandwidth	Max: 20 MHz	Max: 100 MHz (< 6 GHz) Max: 1 GHz (>6GHz)
Carrier Aggregation	Up to 32 Carriers	Up to 16 Carriers
Digital Beamforming	Up to 8 Layers	Up to 12 Layers
Channel Coding	Data: Turbo Coding Control: Convolutional Coding	Data: LDPC Coding Control: Polar Coding
Subcarrier Spacing	15 kHz	15, 30, 60, 120, 240 kHz
Self-Contained Subframe	Not Supported	Supported
Spectrum Occupancy	90 % of Carrier BW	Up to 98 % of Carrier BW

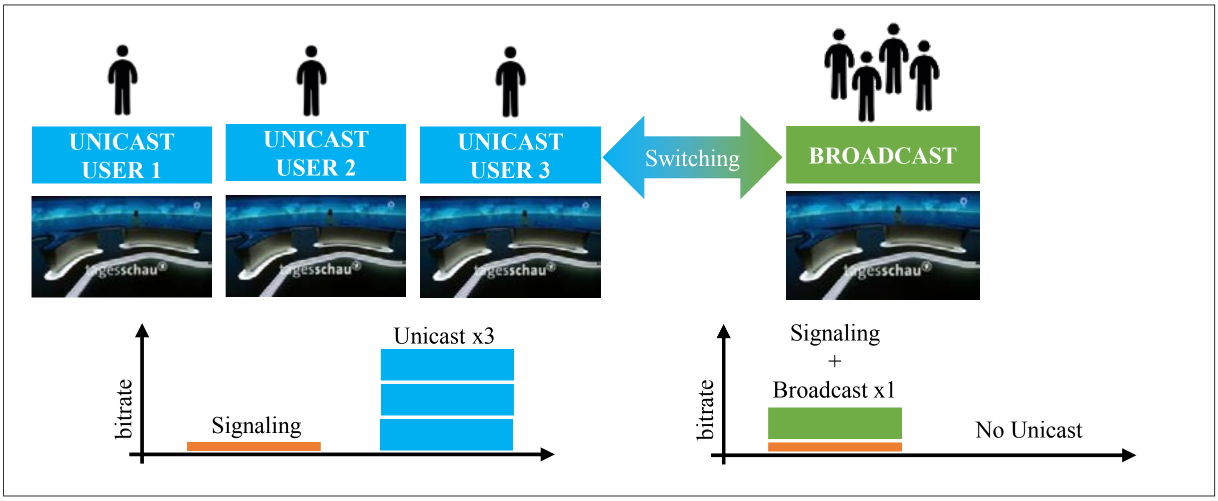

Figure 3. Simplification of the Dynamic switching between unicast and broadcast according to increasing audience demand

and between cells configured with eMBMS or not guaranteeing that the service is not interrupted (seamless transition). MooD then intelligently activates or deactivates eMBMS, ensuring the mobile network is used as efficient as possible in that location.

However, the way eMBMS works may not be the best for certain applications and here is where 5G has an opportunity to break with the existing mechanisms in LTE where broadcast transmissions were initially conceived as a pre-positioned service rather than a mechanism to create a flexible and dynamic way for optimizing system and network performance. One of the main issues with eMBMS is the need of a specific and dedicated system architecture already with little commonalities with the regular unicast architecture. The lack of commonalities with unicast makes it necessary to implement specific functions in the network as well as the user device. And this gets indeed complicated when perhaps the business case is not obvious.

Keeping as many commonalities as possible with unicast may maximize the compatibility and interoperability of unicast, multicast and broadcast delivery mechanisms. Even more, a common architecture may also simplify the way broadcast is used, making it transparent to content providers and users. A common framework fosters that network and user equipment support multicast and broadcast without requiring specific hardware, which is identified as one of the possible barriers to make eMBMS successful worldwide.

The way forward

3GPP should find the way to make the recent 5G developments, the 5G-MSA or even new features such as network slicing fully compatible with a framework where unicast, multicast and broadcast are enabled for an agile and dynamic scheduling of the radio resources. Ideally the use of unicast, multicast or broadcast could be decided dynamically on a cell by cell basis or with the possibility of synchronizing groups of adjacent cells in, at least, small-scale single frequency networks (SFNs). This, without preventing a future evolution to other options like an update of the LTE-based 5G Broadcast technology towards 5G-NR as already designed in [10]. The involvement of all relevant verticals (media and entertainment, automotive, IoT or public warning) will be required so different delivery mechanisms in 5G can provide better spectral efficiency, coverage, less energy consumption, and better latency. Even more, 5G is an opportunity to prioritize solutions which minimize implementation complexity thus fostering a universal adoption of the technology in massive markers.

For the moment, 3GPP is active in standardizing new mechanisms for multicast and broadcast in 5G Release 17 (see [8] and [9]) which may provide solutions to the lack of scalability of current networks. It is now time for vertical industries, including the media sector, to decide whether these new features are interesting or not to deliver their IP-connected service offering.

References
[1] http://www.ard-zdf-onlinestudie.de/ardzdf-onlinestudie-2019
[2] J. J. Gimenez, et at, "5G Broadcast Release 16 is Approaching – Completing a new step to deliver public service media with 3GPP technology", FKT, vol. 1-2/2020, pp. 35-38, Jan/Feb 2020.
[3] 3GPP TS 22.101: "Service aspects; Service principles", clause 32.
[4] 3GPP TR 26.917: "Multimedia Broadcast Multicast Services (MBMS) and Packet-switched Streaming Service (PSS) enhancements to support television services".
[5] 3GPP TR 36.976 v16.0.0, "Overall description of LTE-based 5G broadcast (Release 16)", Dec. 2019.
[6] F. Gabin, et at, "5G Media Streaming Architecture," Proc. International Broadcasting Convention (IBC), Amsterdam, The Netherlands, 2019.
[7] 3GPP TS 26.501: "5G Media Streaming (5GMS); General description and architecture".
[8] 3GPP SP-190625, "FS_5MBS: Architectural enhancements for 5G multicast-broadcast services", SA#85, Newport Beach, United States, Sept. 2019
[9] 3GPP RP-193248, "New Work Item on NR support of Multicast and Broadcast Services", RAN#86, Sitges, Spain, Dec. 2019
[10] J. J. Gimenez, et al, "5G New Radio for Terrestrial Broadcast: A Forward-Looking Approach for NR-MBMS," IEEE Transactions on Broadcasting, vol. 65, no. 2, pp. 356-368, June 2019.


Source: IRT

DR. JORDI J. GIMENEZ

is Project Leader and Research Engineer at the Future Networks Division of the Institut für Rundfunktechnik GmbH
www.irt.de