VoLTE: What makes voice over IP “carrier grade"?

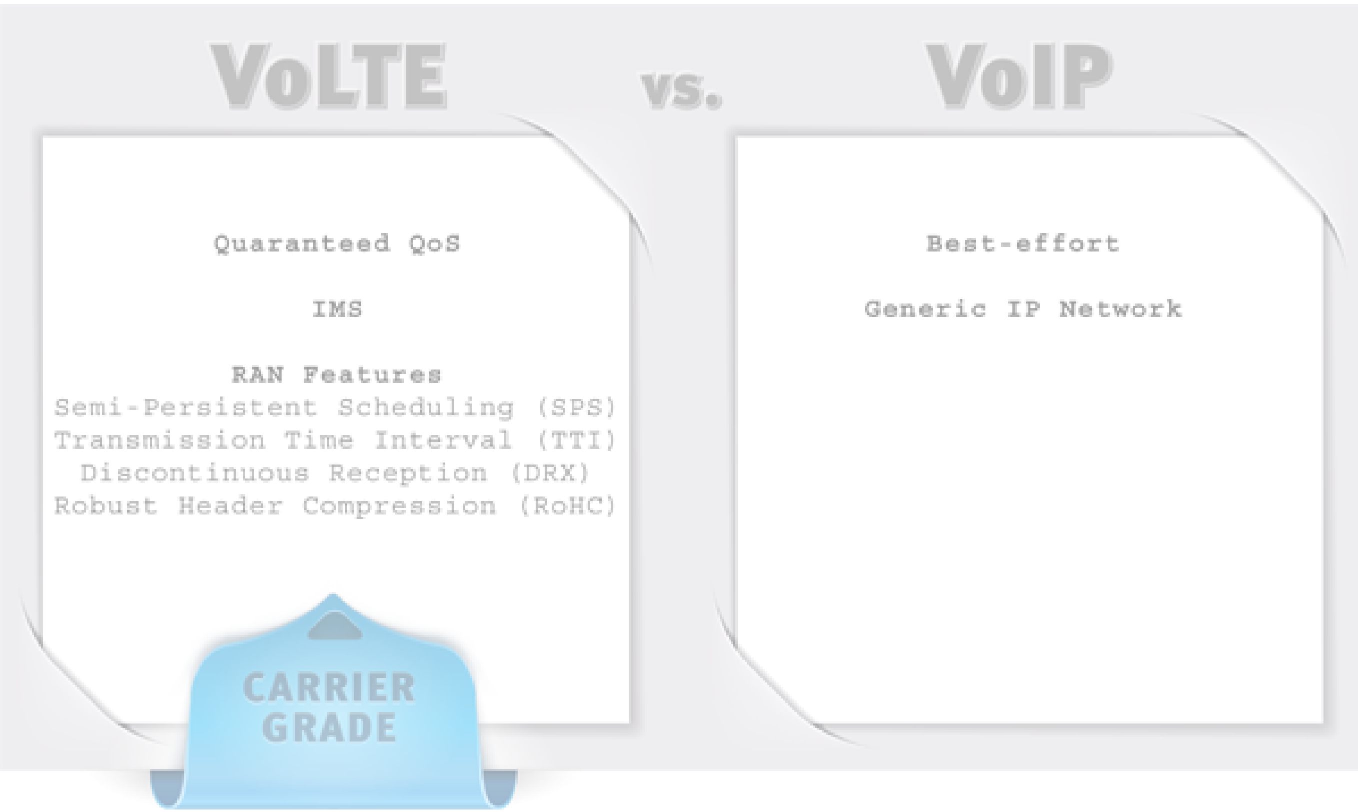
Mike Keely, Director, Product Management, Wireless at Spirent Communications

At first glance Voice over LTE (VoLTE), the next generation of wireless voice technology, looks suspiciously like VoIP service over a wireless connection. At a very basic level it is exactly that: digitized voice-band audio transmitted as IP packets and demodulated into voice-band audio. Then what is the difference between VoLTE and over-the-top services like Skype, Vonage or Gizmo5?

The driver for VoLTE is the expectation for“carrier-grade” (or “telco-grade”) voice services that are at least as good as, if not better than, legacy circuit-switched voice services when perceptually measured by the customers.This concept fundamentally differentiates VoLTE from VoIP.

The ubiquity of IP led to its selection as the de facto vehicle for all digital data, including voice. However, generic IP still involves best-effort delivery, meaning that latency cannot be guaranteed. For many of us though, voice service is a critical part of our daily lives. Although some OTT VoIP services currently work reasonably well, we can’t safely assume that the networks of tomorrow will always have enough available capacity to rely on best-effort service. The solution is VoLTE.

Implementing VoLTE



The IMS Subsystem
One “headline” aspect of implementing VoLTE is that it is deployed on the IMS

subsystem rather than a generic IP network (i.e. the Internet). This offers a level of control that is not available otherwise.

The relationship between VoLTE and IMS was a pretty significant factor when the industry was deciding how it would implement digital voice. Proponents of competing technologies (e.g. Voice over LTE via Generic Access, or VoLGA) were fond of pointing out that VoLTE required the costlyimplementation of IMS. However, since many network operators had other reasons to deploy IMS, this was treated as a sunk cost, making the business case for VoLTEas the industry's choice for carrier-grade digital voice.

QoS
One primary driver for the development of VoLTE is the concept of Quality-of-Service (QoS). This generic-sounding phrase has a specific meaning in communications. QoS describes the delivery of a minimum level of latency, availability and data integrity. Sometimes QoS is defined on a per-user basis (such as a service level guaranteed by a premium data package) and sometimes on a per-application basis (such as VoLTE).

For network operators to deliver specific QoS levels, they need to have control over every link in the end-to-end service chain. This is impossible when depending on the uncertainties and best-effort delivery mechanisms associated with the Internet. Delivering VoLTE via the IMS subsystem lets operators add value by committing to QoS levels.

VoLTE from the UE's point of view
The inner workings of VoLTE are defined in a GSM Association (GSMA)document. While its title is IMS Profile for Voice and SMS, industry insiders simply refer to its document number, IR.92.

IR.92 is a profile of minimum IMS capabilities and services needed for telephony, including real-time media negotiation, transport, codecs, LTE radio and Evolved Packet Core (EPC). The document presents a high-level view of what should happen. The details that define how things happen are not so simple; note that IR.92 refers to 12 IETF documents and 46 separate 3GPP specifications!

Technical requirements for supporting VoLTE
From the UE's point of view of the network, there are four non-obvious network requirements for supportingVoLTE. The first three of these, Semi-Persistent Scheduling, Transmission Time Interval Bundling and Discontinuous Reception, are implemented at the MAC sub-layer. The fourth, Robust Header Compression, is implemented in the Packet Data Convergence Protocol (PDCP) sub-layer.

Semi-persistent scheduling (SPS)
In most LTE data usage cases, every Physical Resource Block (PRB) on the downlink and uplink must be explicitly granted; the resulting overhead is inefficient for traffic that requires continual allocations of small packets (such asVoLTE).

This issue is addressed by SPS, which defines a transmission pattern and, based on that pattern, assignsa pattern for PRBs to use going forward (unless there is a reason to change the pattern). As an example, suppose a voice service uses one coded packet every 20ms. During silent periods, PRB assignments can be canceled. In the uplink they can be implicitly canceled after a defined number of empty UL transmissions. In the downlink they can be canceled with a Radio Resource Control(RRC) message.

Transmission time interval (TTI) bundling
In order to reduce end-to-end latency, LTE introduced the idea of the short TTI (1 ms). This means that the Hybrid Automated Request (HARQ) process is meant to acknowledge transmissions every 1 ms. However, at cell edges, a UE might not have enough time available to reliably deliver an entire VoIP packet in one TTI.

The solution is to bundle multiple TTIs together without waiting for HARQ feedback. A VoIP packet is sent as a single packet data unit (PDU) during a bundle of subsequent TTIs, and the HARQ feedback is only expected after the last transmission of the bundle. As in legacy technologies, RRC protocol is used to configure TTI bundles.

Discontinuous reception (DRX)
A constantly-on voice session can quickly reduce battery life. Since VoLTE traffic is predictable (e.g. 20ms codec packets), a UE receiver does not have to constantly monitor the physical-layer control channel, and the receiver can essentially be turned off between receptions. This must be carefully configured, though, since missing acknowledgements or HARQ messages can add unacceptable latency.

Robust header compression (RoHC)
IP header information can be disproportionately large when compared to the relatively small VoLTEcodec packets being transmitted, creating inefficiency in terms of the air interface bandwidth.

For example, a combination of RTP, UDP and IP headers can total 40 to 60 bytes of header data, while using the AMR-WB codec at 14.4 kbpsyields payload data of about 50 bytes per 20 msframe. In this case there may be more overhead being transmitted than actual payload data. RoHC can sometimes compress headers down to the 2-4 byte range, providing greatly improved efficiencies on the air interface.

Conclusion
Although IMS usually grabs the headlines as being essential for deploying VoIP in an LTE environment, it is ultimately the implementation of the LTE RAN features described in this article that will enable carriers to differentiate clearly between VoLTE and VoIP. While much attention to date has been placed on testing the UE’s IMS connectivity, it is fully-integrated testing of a UE’s signaling along with the negotiation, establishment and usage of the previously discussed RAN features that will be one of the most critical hurdles in successful deployment of carrier-grade VoLTE.

About the author

Michael Keeley is a Director of Product Management at Spirent Communications’ wireless test equipment division. He has led various teams involved in wireless network emulation and automated systems used for testing mobile devices. Prior to joining Spirent in 2000, Mike worked for Lucent Technologies. He earned his BSEE and MEng from Cornell University and an MBA from New York University’s Stern School of Business.

Source URL (retrieved on *03/29/2015 - 9:14am*):
http://www.ecnmag.com/articles/2012/09/volte-what-makes-voice-over-ip-%E2%80%9Ccarrier-grade?qt-recent_content=0&qt-most_popular=0