Loquendo MRCP Server: the importance of being standard

Roberto Pacifici, Loquendo

1. Overview

This article will focus on the Loquendo MRCP Server (LMS), a 'speech server' (i.e. the base speech technology functionalities are made available through a suitable integration layer for multichannel environments), and the advantages provided to System Integrators by the adoption of MRCP standards.

The opportunity to release a 'speech server' onto the market occurred when certain standards became widely adopted by the major commercially available platforms: this refers particularly to the use of the MRCP protocol, both MRCPv1 [1], the version most established on the market, and MRCPv2 [2,3], currently gaining ground. The use of standards brings a great many benefits, in particular the choice to adopt standards, it is worth noting, supports integration between clients and servers from different vendors (*). Commercially available MRCP platforms (MRCP clients) can be divided into two groups: platforms that natively support the MRCP protocol, normally MRCPv1, and platforms to which MRCP support is added at a later time, when the need arises to make use of speech technology to improve the quality of services offered to the end user: in this case MRCPv2 is often chosen, being the latest version of the protocol and more powerful in several different ways, as we will discuss in more detail later in the article.
In the pages that follow we will explore the Loquendo solution more closely, and look at some of the benefits it has to offer.

2. The Loquendo MRCP Server

While speech technologies (Text To Speech and Automatic Speech Recognition) have traditionally been integrated via proprietary APIs, the new method now gaining ground is to rely on MRCP protocol to access a 'speech server', more suitable for those environments making use of a client-server architecture: typically where large scale distributed speech technologies are required, such as Automated Call Distribution (ACD), Contact Centres, etc. In such contexts, the use of a 'speech server' makes it possible to create personalised services with a significant cost reduction compared to analogous solutions using live agents.
Laid out in Figure 1 is a schematic representation of Loquendo's speech technology products, showing the different interfaces offered to integrators.

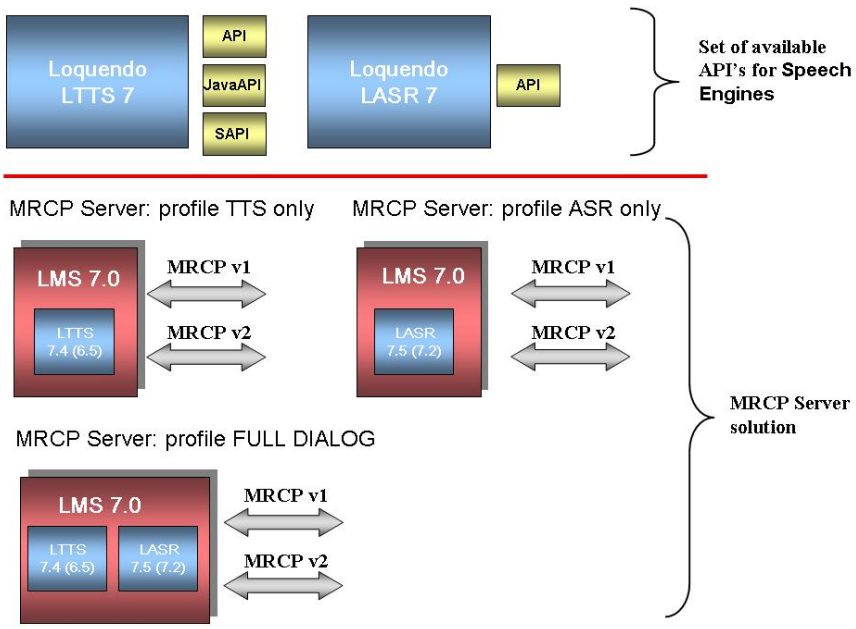

Figure 1: Loquendo speech technology integration scenarios

Loquendo's MRCP Server (LMS) can be configured for MRCPv1 or MRCPv2, thus responding to current, differing market demands. As noted above, MRCPv2 varies significantly from MRCPv1, both in the way it establishes control of the MRCP session (based on the SIP protocol [4]), and also in that MRCPv2, as well as managing speech

synthesis and speech recognition resources (as does MRCPv1), also manages Recorder and Speaker Verification resources, available with Loquendo technologies. Independently from the MRCP protocol chosen, LMS can be configured for speech synthesis (TTS only profile), for speech recognition, recorder and speaker verification (ASR only profile), or for "Full Dialog" profile, using TTS and ASR technologies together, as shown in Figure 1.
Another aspect that should be considered is the ease with which LMS can be connected to a pre-existing, pre-configured LAN network, which can be done by installing it on an independent server making it accessible to several MRCP clients, all simultaneously present on the LAN network.

The LMS architecture is shown in Figure 2, indicating the standard product interfaces and the internal interface used by the MRCP layer to dialog with the 'technology layer'. From the drawing it can be seen that LMS uses Loquendo TTS 7 and/or Loquendo ASR 7 as a technology layer, interfacing with them by means of LTTS and/or LASR standard APIs - to which it adds a series of software modules to manage the multichannel capability as well as the communication protocols. The clearest advantage of this architecture derives from the separation of the MRCP server layer and the technology layer, allowing the rapid integration of new versions of Loquendo technologies into the server, and the easy upgrade to the latest protocol evolutions (typically for MRCPv2).
It is worth mentioning that Loquendo technologies offer features that, even if not directly covered by the MRCP protocols, enable the MRCP client with which the LMS is connected to make use of additional capabilities that can be extremely useful: these can be invoked by means of the 'vendor specific' activation parameters covered by both MRCP protocols and described in detail in the product documentation.

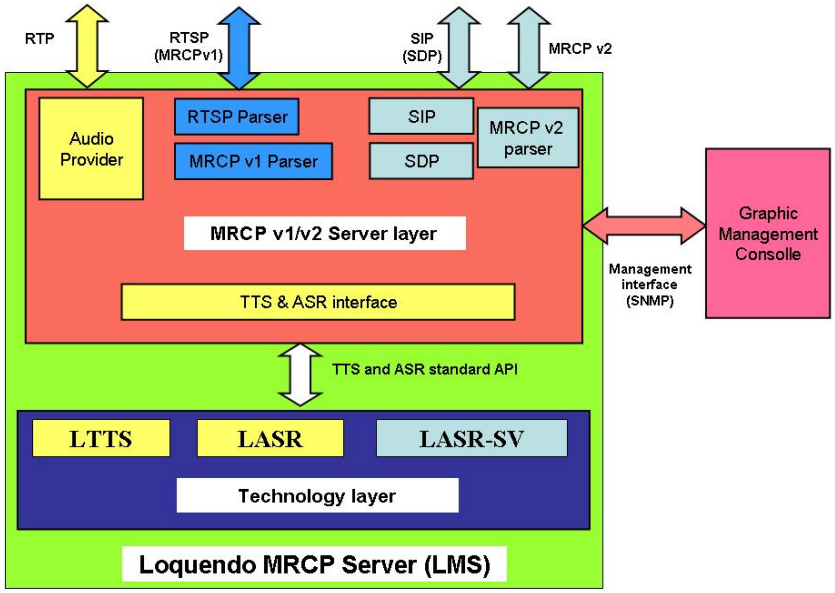

Figure 2: LMS Architecture

Among the standard interfaces managed by the LMS there is also the 'Management Interface', based on the SNMP protocol [5], that allows, by means of the 'Management Console' graphical tool, the administration, configuration and monitoring of the LMS: the 'Management Console' software can be installed on the same server as the LMS, but it can also be installed on a remote PC, enabling the management of several LMS servers from a single console; By means of the SNMP protocol [5], the LMS can also be managed by a third-party commercial management tool other than the one provided by Loquendo.

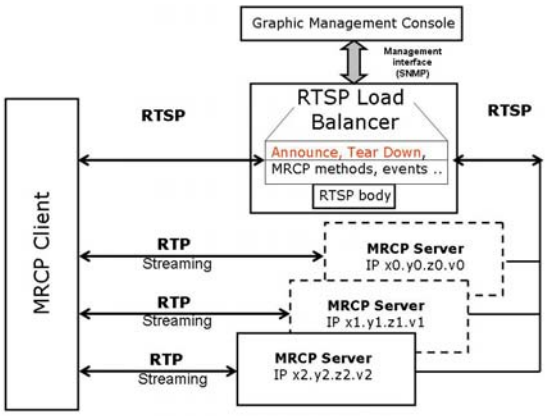

Figure 3: RTS Load Balancer for LMS

The Loquendo MRCP Server also includes an optional module: the Load Balancer. This has been designed to meet the needs of a single MRCP client that interoperates with several Loquendo MRCP servers (all configured according to the same profile but each, let's imagine, with a different maximum number of sessions available: this could be, for example, because the servers have different CPU/RAM capabilities).
With this configuration, the MRCP Client module connects directly to the Load Balancer rather than to the MRCP Server, so it is the Load Balancer itself that

redirects the requests (made by the MRCP Client) to one of the servers under its control, to open new MRCP sessions, according to the number of sessions active at that time on each single MRCP Server: if the MRCP Servers associated with the Load Balancer have different CPU/RAM capabilities, the Load Balancer will balance out the load among all the MRCP servers it controls.

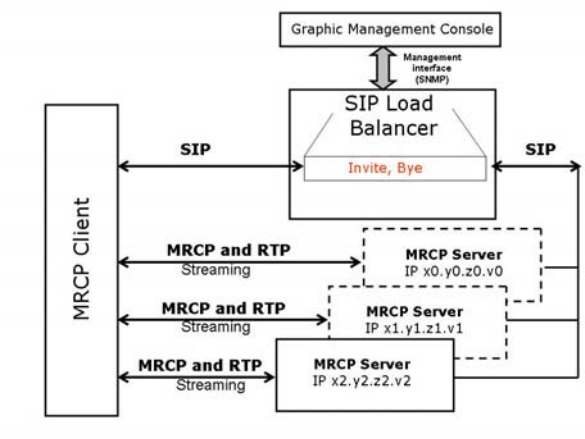

Figure 4: SIP Load Balancer architecture

The introduction of the Load Balancer module enables an increase in the scalability of MRCP sessions controlled by a single MRCP Client, as well as improving the robustness of the system; in fact, if one or more LMS becomes unreachable (for network problems, for example), the new MRCP sessions requested will be controlled by another LMS from the still active pool.

The Load Balancer can also be administered, configured and monitored by means of the Management Console: during the configuration process, in fact, it is possible to define a 'computational weight' for each single MRCP server that will assist the Load Balancer in correctly balancing the load.

Two versions of the Load Balancer are made available, due to the fact that the MRCP protocols supported are significantly different: one for MRCPv1 (RTSP Load Balancer) and one for MRCPv2 (SIP Load balancer). The architectural logic that enables the use of the Load Balancer between MRCP client and MRCP server is shown in Figure 3 (for MRCPv1) and Figure 4 (for MRCPv2).

3. Loquendo MRCP Server: a Summary of the major benefits

In conclusion, it is in many ways thanks to the company's adoption of standards that Loquendo has managed to successfully deliver the Loquendo MRCP Server to the market – offering a client-server solution optimized for the large-scale deployment of speech technologies in many different fields: Finance, Corporate, Public Sector, Telco/Internet, Call Centres, CRM, messaging, self-service applications, etc. Such deployments are highly cost-effective thanks to Loquendo MRCP Server's design, which enables speech resources to be shared between different yet simultaneous calls.

System Integrators looking for high-quality multilingual speech technologies, and looking to boost the performance of existing IVR services using speech, can now more effectively exploit Loquendo ASR and Loquendo TTS using a speech server configurable as TTS-only, ASR-only, or both TTS and ASR; moreover, because of the flexibility of the LMS, this is possible without the need for specific, speech-related technical expertise. For those willing to finely adapt the technology to their specific needs, on the other hand, a set of powerful tools are available(**) enabling Loquendo partners to fine tune speech engine parameters and improve speech applications, without the need for costly, professional services.

(*)Loquendo MRCP Server is certified with the products of major IVR Vendors such as Avaya, Cisco, Genesys, Interactive Intelligence, Envox, Siemens, Aculab, NMS, Voxpilot, App-line and others.

(**) Loquendo TTS SDK and Loquendo ASR SDK are delivered separately, on request.

References

[1] “A Media Resource Control Protocol (MRCP)”, RFC-4463 Informational, IETF, http://www.ietf.org/rfc/rfc4463.txt

[2] “Media Resource Control Protocol Version 2 (MRCPv2)” Internet Draft, IETF, http://ietfreport.isoc.org/all-ids/draft-ietf-speechsc-mrcpv2-09.txt

[3] Dave Burke, “Speech Processing in IP Network. Media Resource Control Protocol (MRCP)”, ed. John Wiley & Sons, 2007.

[4] “Session Initiation Protocol (SIP)”, RFC3261, IETF, http://www.ietf.org/rfc/rfc3261.txt

[5] “Simple Network Management Protocol (SNMP)”, RFC 1157, IETF, http://www.faqs.org/rfcs/rfc1157.html

The Author

Roberto Pacifici - Roberto joined CSELT (Telecom Italia’s Research Center) in 1974. His principal working experience has been in Broad Band Amplifiers and Speech and Signal processing, where he was involved in the porting of speech recognition and echo-cancelling algorithms to DSP boards. He has published several papers in these areas. He joined Loquendo in 2001, where he is currently MRCP Server Product Manager.