



# “SpeD 2009”

June 18 – 21, 2009 • Constanta, ROMANIA

---

## “Real-Time Architecture For A Network- Based Text-To-Speech Service Implementation”

*Mihai Surmei<sup>\*</sup>, Dragos Burileanu<sup>\*\*</sup>, Cristian Negrescu<sup>\*\*</sup>,  
Catalin Ungurean<sup>\*\*</sup>, Aurelian Dervis<sup>\*\*</sup>*

**\* ERICSSON Telecommunications Romania S.R.L.**

**\*\* Faculty of Electronics, Telecommunications and IT,  
University “Politehnica” of Bucharest, ROMANIA**



# OUTLINE

---

- **Goals**
- **Multi-service network architectures**
- **TTS network services**
- **A real-time processing environment for speech synthesis in Romanian**
- **TTS in IMS context – reference architecture**
- **TTS engine**
- **The proposed service: PoC-Chat**
- **Conclusions**



# Goals

---

- **Build up a reference network-based media processing environment for Romanian TTS**
- **Particular service on proposed reference environment combining several network capabilities**



# Multi-service network architecture

---

- From peer-to-peer voice and shoot-and-forget messaging to session oriented real-time multi-service communication
  - *Mobile networks*
  - *Fixed networks*
  - *Nomadic networks*
  - *Internet*
- The common factor: IMS

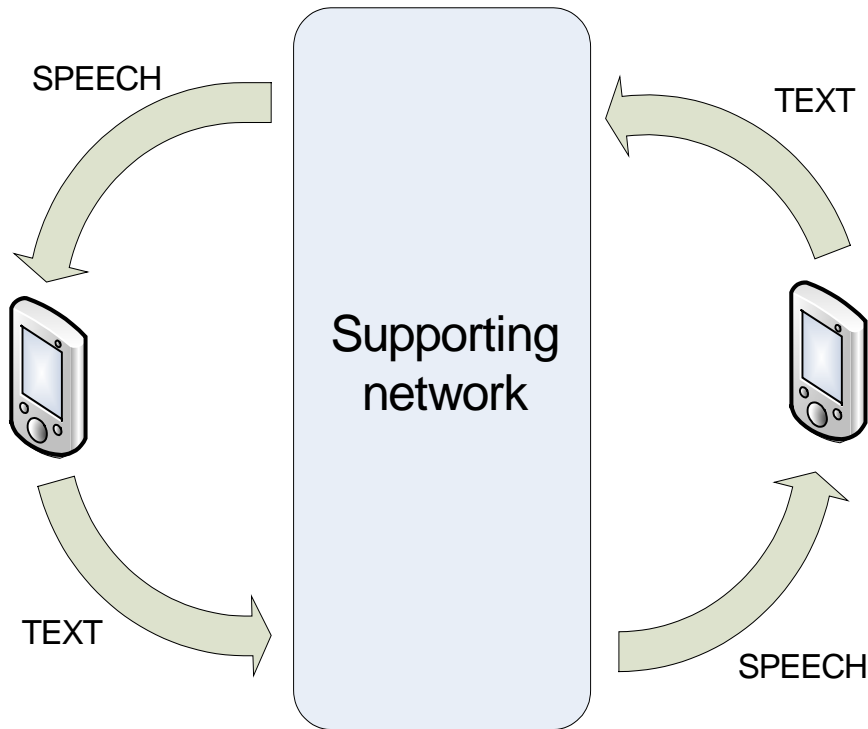


# TTS network services

---

- **Existing TTS services**
  - **Not real-time**
  - **Proprietary client-server approach**
- **Next generation services**
  - **Real-time**
  - **Service mix**
  - **Open architecture and protocols**

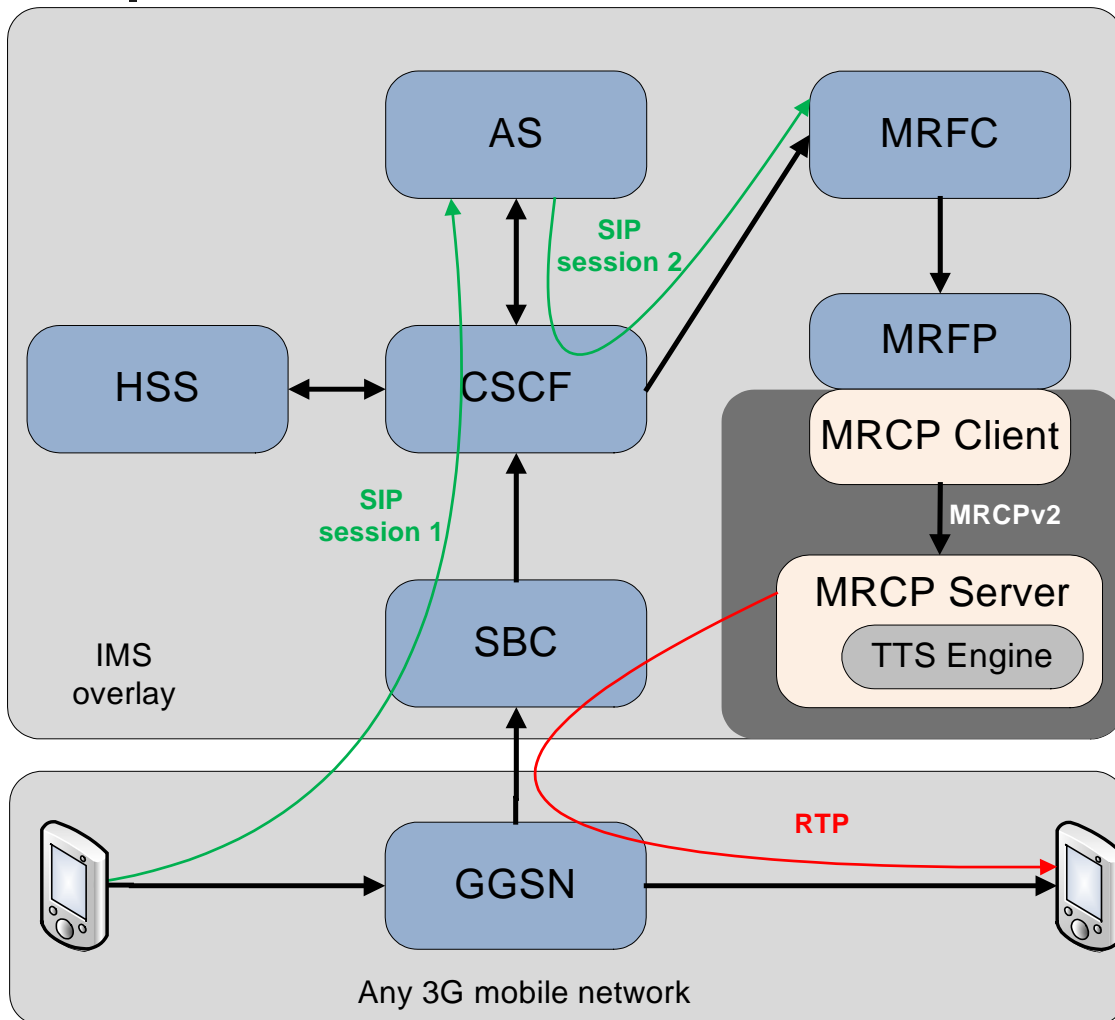
# A real-time processing environment for speech synthesis in Romanian



Generic real-time TTS-based telecom service

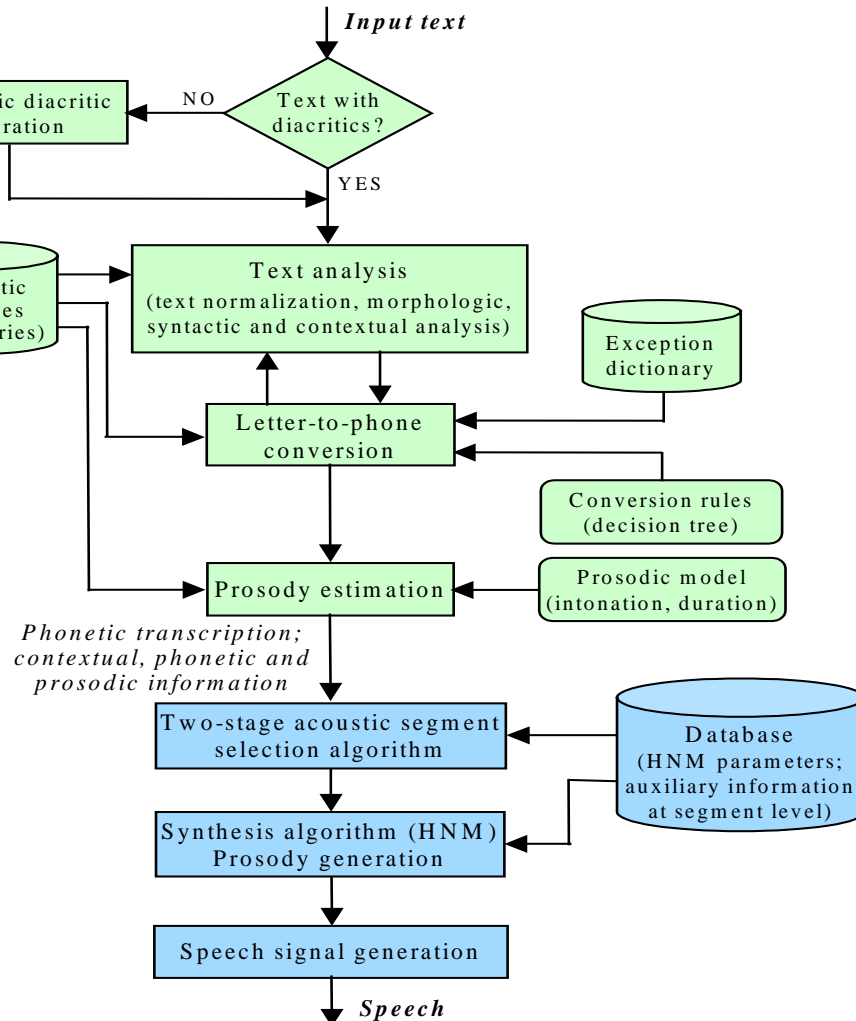
- Leveraging on open protocols (MRCP)
- Following the latest development in telecom field
- Modular design
- Expandable to close the loop on speech recognition

# TTS in IMS context – reference architecture



- **IMS overlay network: supplementary control**
- **Access agnostic: GPRS/HSPA, ADSL, WiFi**
- **TTS closely related to MRFC/MRFP pair due to the hybrid functionality:**
  - **Signaling**
  - **Payload**
- **MRCP protocol**

# TTS engine



- The last version of our concatenative TTS system is based on non-uniform acoustic units (diphones and polyphones). The synthesis technique makes use of the Harmonic plus Noise Model of speech.
- The TTS system implementation has been enhanced in order to be integrated in a multi-thread/multi-process environment.

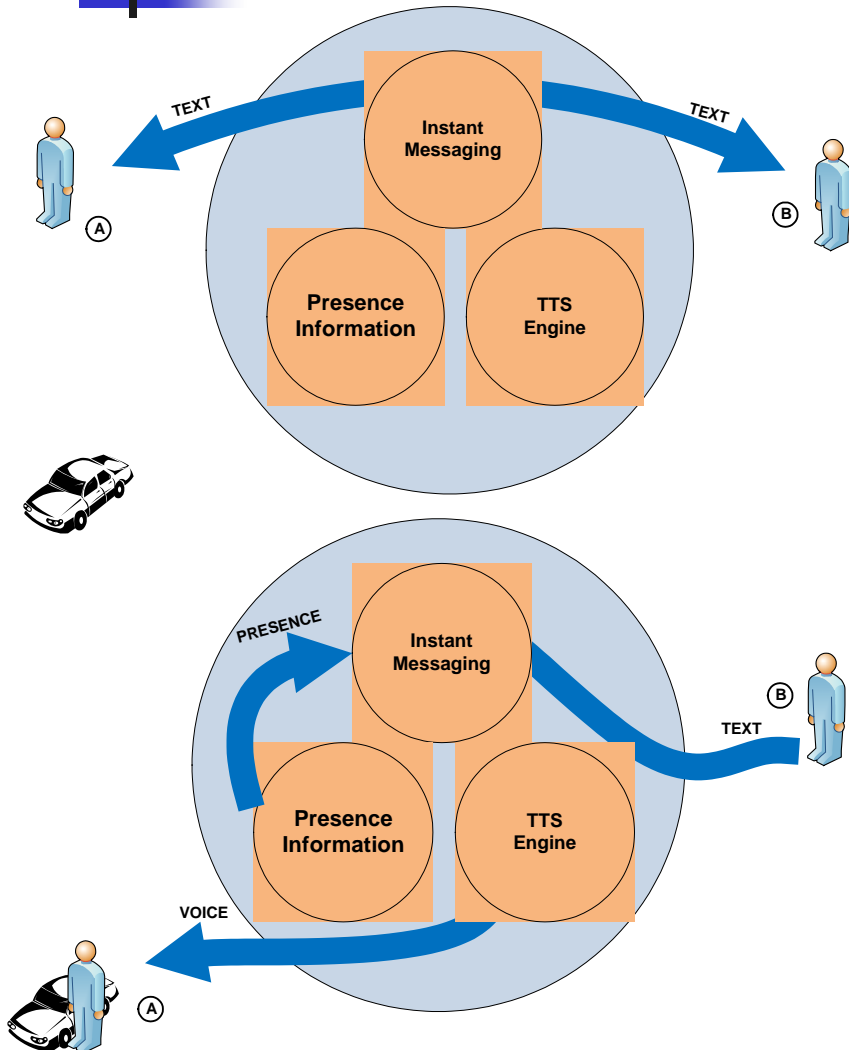


# The proposed service: PoC-Chat (1)

---

- **Paradigm shift towards multi-session convergent experience**
- **IMS network exposes services**
- **Our proposal - a service mix:**
  - **Text (chat)**
  - **Voice**
  - **Presence**
  - **TTS conversion**

# The proposed service: PoC-Chat (2)



- A and B users are chatting
- A change state, network reacts switching on TTS conversion
- A “hears” the chat conversation



# The proposed service: PoC-Chat (3)

---

- **IMS test network using existing functional nodes: CSCF, HSS, AS**
- **ACE framework for MRCP server:**
  - **TTS engine**
  - **RTP stack**
- **SIP servlet technology for MRCP stack**



# Conclusions

---

- **To emphasize the TTS importance, we add it into the larger set of existing telecom services**
- **We presented a convergent service example combining TTS, IM and presence**
- **IMS overlay network will allow mixing the needed capabilities on a single session**
- **Service realization is based on a modified TTS engine, new MRCP server and client developments**