3GPP T2, TSG T2 (01)0644
Background for T2-010627
Agenda X.X

DSR in 3GPP

Contact:
Stéphane H. Maes,IBM
smaes@us.ibm.com
for ETSI STQ Aurora

Edinburgh - September 2 - 7 , 2001

SOURCES

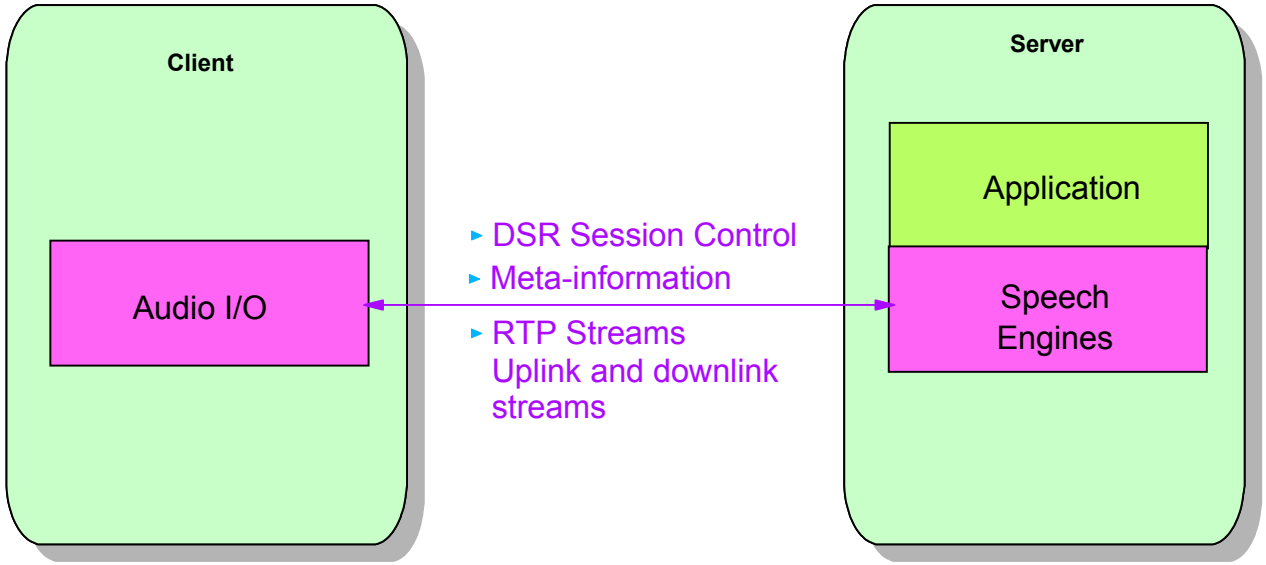
- 3GPP:
  - Alcatel, France Telecom, IBM, Mannesman, Motorola, Omnitel, Qualcomm, Texas Instruments, Vodafone
- ETSI STQ Aurora:
  - Standardize feature extraction (DSR Codec)
    - WI007 - Default Codec Standardized as ETSI Standard ES 201 108 for DSR
  - Standardize DSR Protocol Stack
    - IETF AV RTP DSR: RTP Payload format
    - DSR Session Control (SIP) and Meta-information (Work in progress)
  - Target a reference Multi-modal architecture (work in progress)
  - Work closely with other standards bodies.
  - Participating companies:
    - Alcatel, Qualcomm, Motorola, France Telecom, Nokia, Texas Instruments, Siemens, British Telecom, IBM, ICSI, HP, SpeechWorks, VerbalTek, Conversay, Ericsson, Temic, Intel, Philips, etc.

DSR Applications

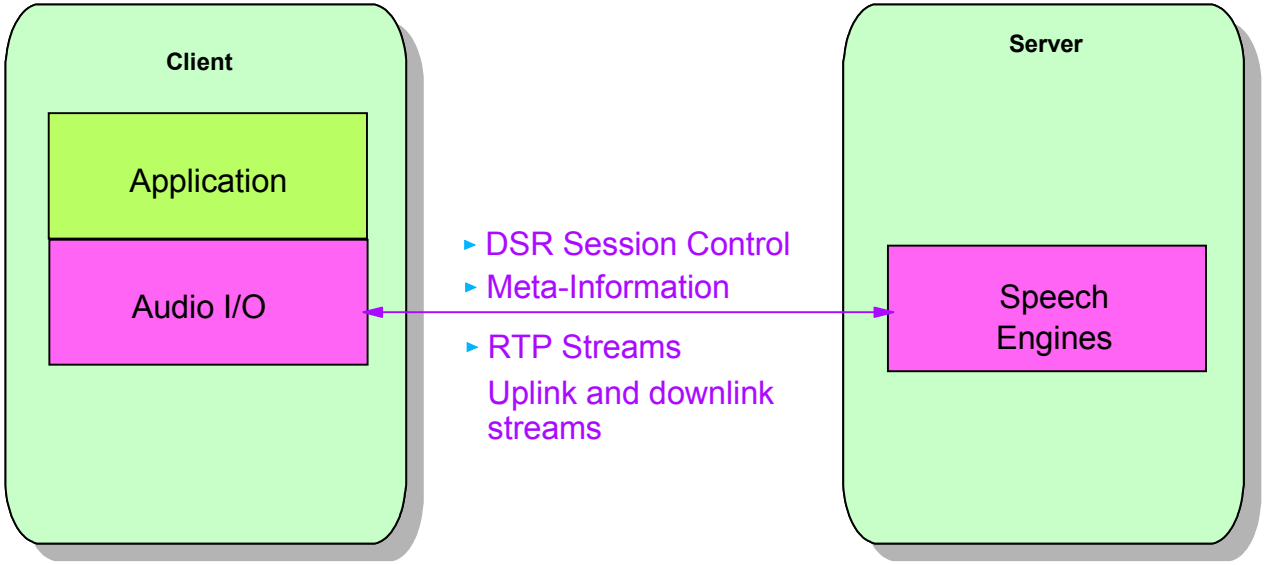
- **Server-Side Voice Recognition Applications:**
  - Voice services:
    - Name dialing
    - Directory Assistance
  - Information applications
    - Voice portals (flight, weather, news, movies)
    - Location-specific information
    - Voice Navigation of maps
  - Transaction-based applications
    - Banking, m-commerce, etc.
    - Form Filling
  - Information capture
    - Dictation
  - Messaging

Usage Scenarios for DSR within 3GPP

Main Envisaged Scenarios: Server-side Applications



Possible Future Scenarios: Client Side applications with remote speech functions



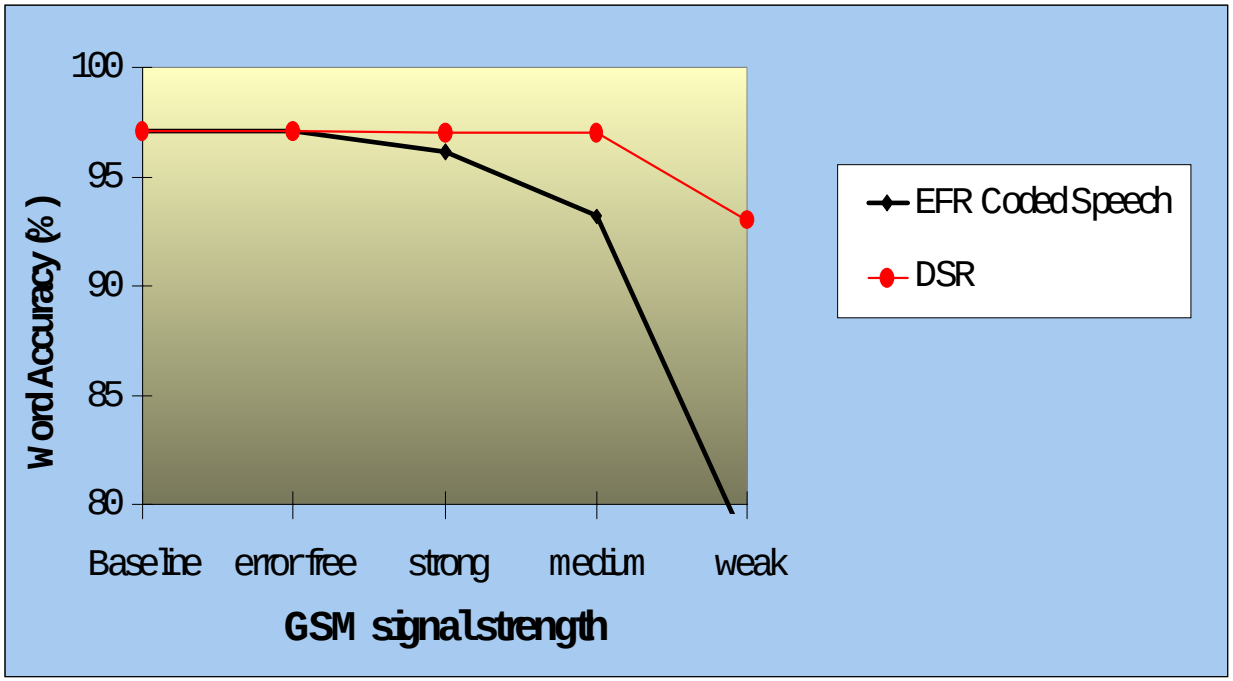
Principles of Distributed Speech Recognition

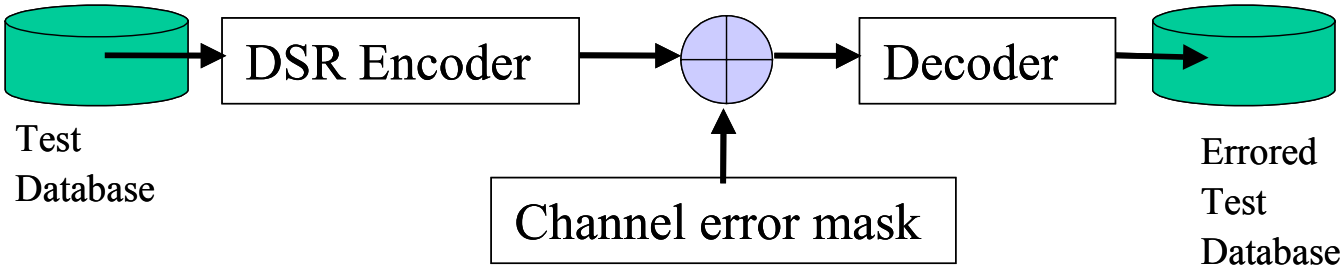
- Distribute the task (DSR)
  - Uplink:
    - DSR frontend (encoder) on terminal.
    - DSR backend (decoder and speech engines) in the network.
  - Downlink:
    - Conventional audio stream (prompt, audio output)
  - Additional uplink and downlink traffic (meta-information) e.g. downlink graphic information
- Benefits of DSR
  - Acoustic frontend features not degraded by codec and network transport
    - Increased robustness and independence from codec, transport and network
  - Speech functions not limited by client capabilities or network bandwidth

Drivers for Distributed Speech Recognition

- With respect to server-side applications with conventional speech exchanges
  - Performance degradations for telephony or wireless speech recognition if transmitted as conventional voice channel
- With respect to client-side applications with speech functions performed on the terminal:
  - Limitation of the client-side speech functions that can be provided on the terminal:
    - Limited client resources with respect to the speech engine requirements (CPU Memory)
    - Too low bandwidth to send data files from the server to a client speech engine or delays introduced on application
    - Proprietary aspect of such data files (grammars, acoustic models etc…)
    - Specialized Speech Engines
- ...

Recognition Accuracy Impact



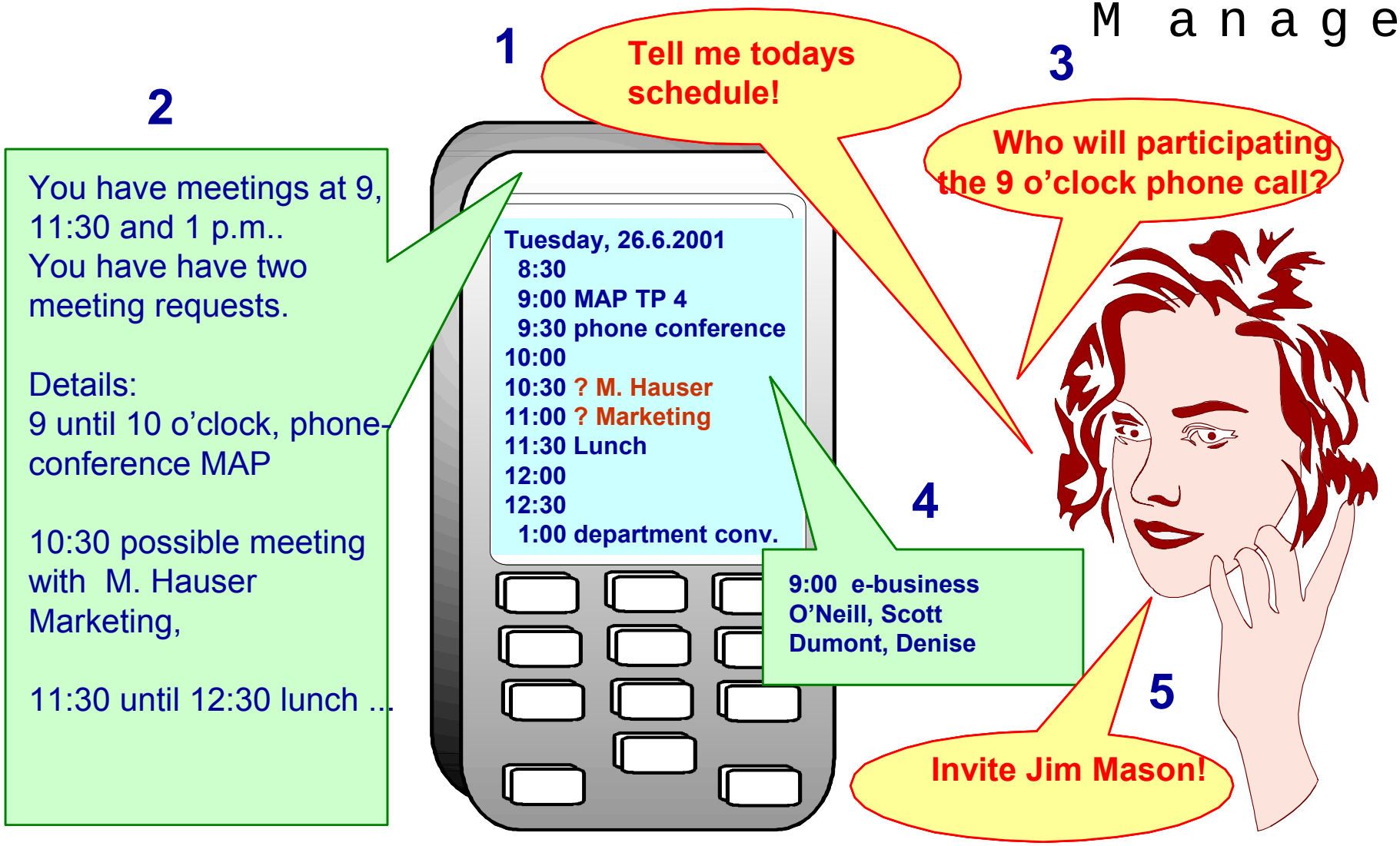


Channel	Performance
Unquantized	97.35
error free quantize	97.45
TETRA TU50 20dB	97.16
GSM EP1	97.44
GSM EP2	97.39
GSM EP3	92.25

- Aurora 2 noisy TI digits task: Connected digits in noise.
- 3GPP signal strength
- GSM Error masks EP1,EP2 and EP3 (outside coverage)
- DSR provides improved accuracy - critical for acceptance of speech applications, especially as perplexity/complexity of tasks increases

Future Usage Target: multi-modal applications

Scenario: Personal Information Manager



See also: T2-010586

IMPACT ON TERMINAL

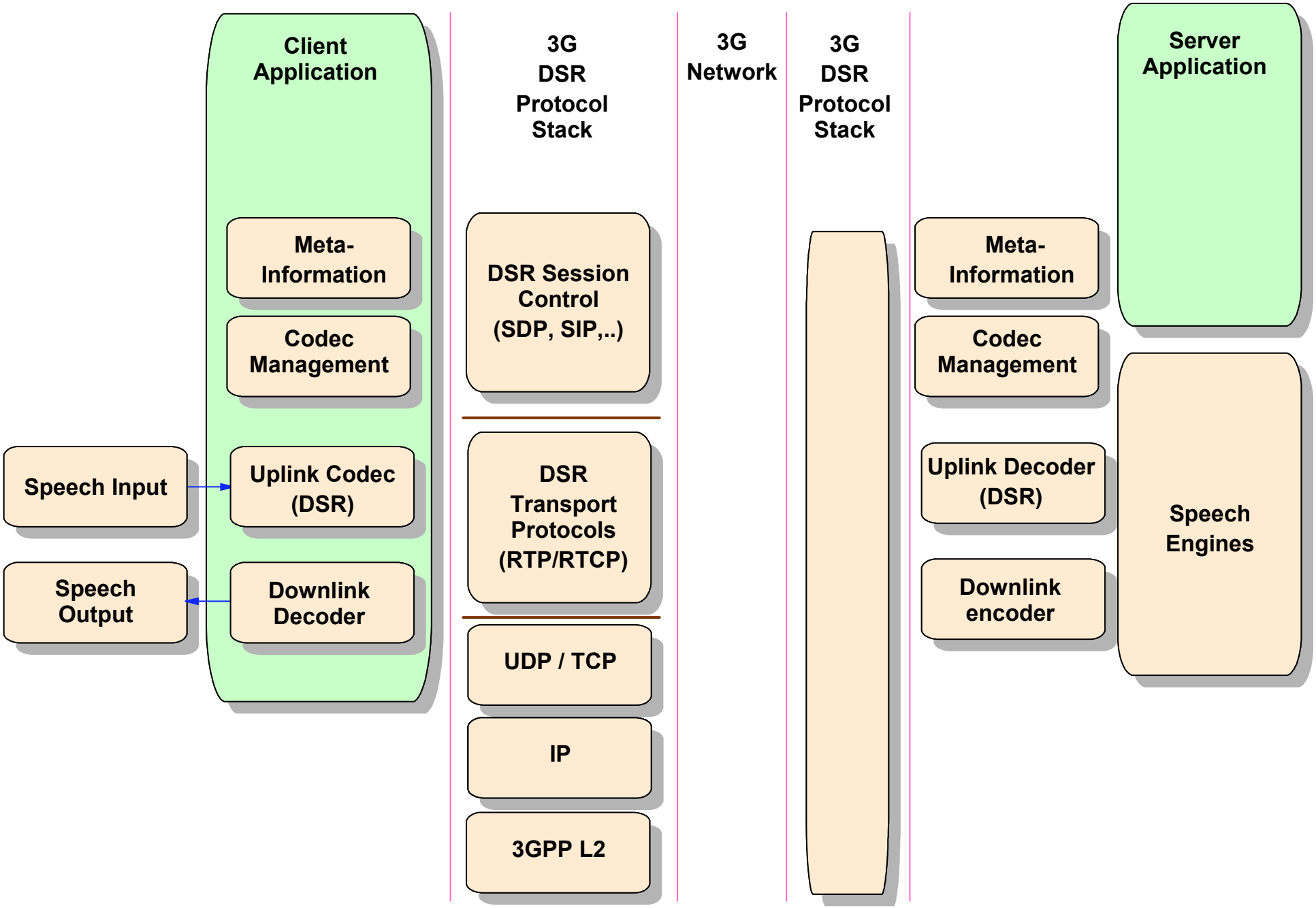
- DSR Codec:
  - Low CPU and memory requirements
    - It can fit within DSP capabilities of client terminal

	wMOPS	ROM(kwords)	RAM(kwords)
Acoustic Front-end	2.5	0.6	0.6
Compression	0.6	1.5	0.03
Total	3.1 (<4MIPS)	2.1	0.6

Memory: Word=16 bits ; Computation: weighted MOPS
By Comparison: AMR Speech Codec ~ 17 wMOPS

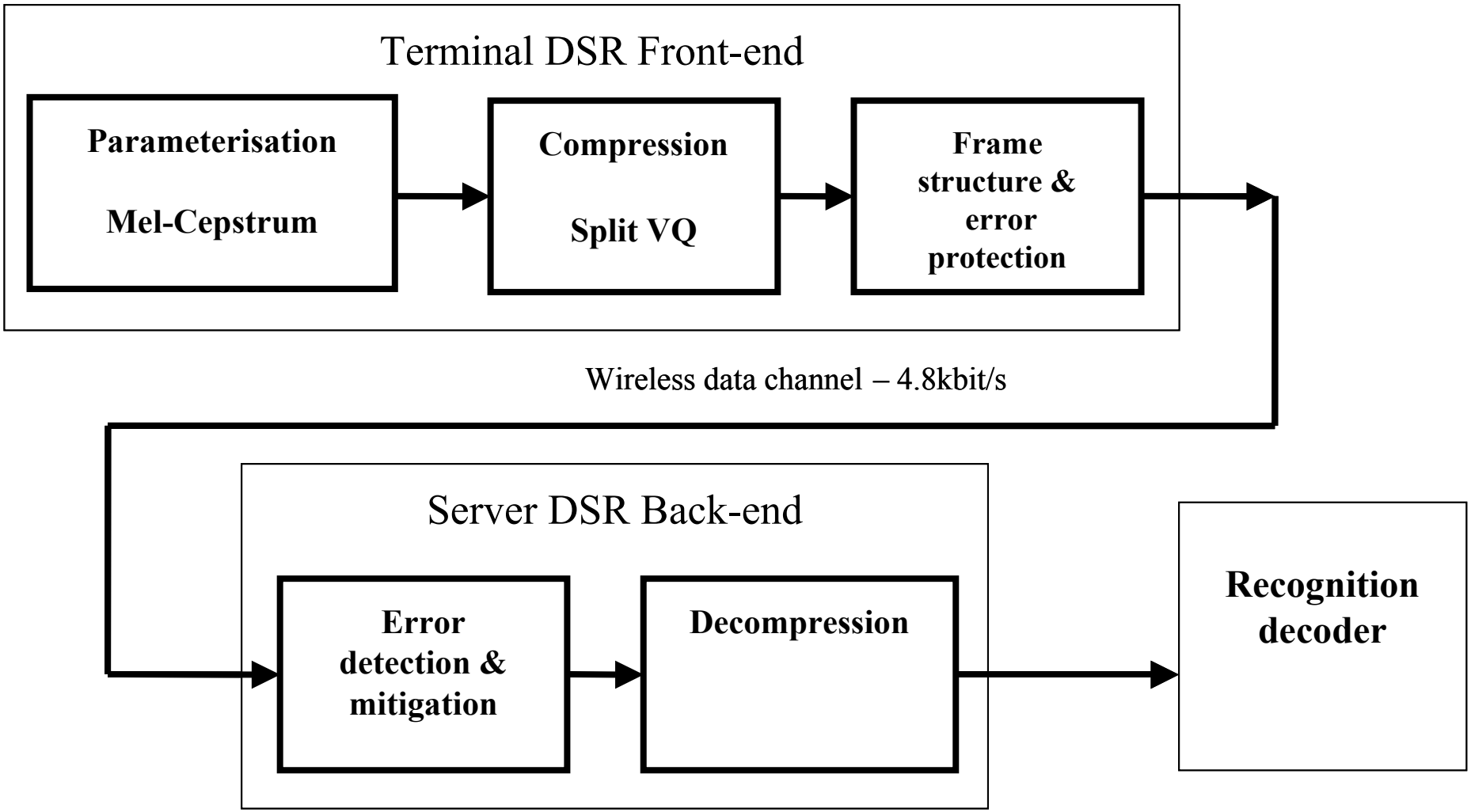
    - Bit rate: raw bit rate from DSR Codec < 4.8 kbit/s
  - Additional terminal requirements:
    - Support for RTP streaming
    - Support for SIP:
      - SDP
      - Some meta-information (e.g. keypad information, ...)

DSR Protocol Stack



CONCLUSIONS

- Speech enabled UI will enhance users' experience in 3G terminals and services.
- ETSI Aurora's DSR initiative. Supported by most handset and ASR vendors.
- Stable ETSI Aurora DSR frontend standard.
- DSR Protocol Stack in IP networks to be completed in time for Release 5.
  - Designed to be compatible with 3GPP framework
- Terminals are capable of supporting DSR requirements

BACKGROUND

ETSI ES 201 108 for DSR