H.VM-VMIA Implementation of Vehicular Multimedia Systems

Media format and control
Smartphone Infotainment assistant

ITU workshop 11 July 2023

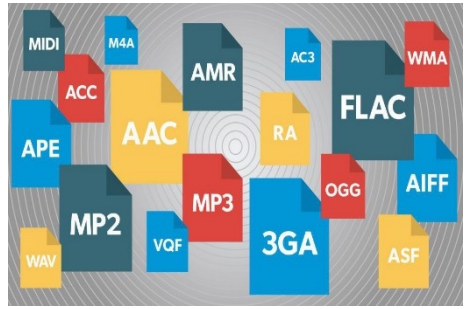
Agenda

- **This presentation aims at showing the current requirements in the WI draft H.VM-VMIA and collecting feedback**

- **Following normative sections are covered in this presentation:**
  - *6.1 Connectivity (Presented by Changan Automobile)*
  - *6.2 HMI (Presented by Changan Automobile)*
  - **6.3 Media format and control**
  - **6.4 Smartphone Infotainment Assistant**

6.4 Media format and control

6.3 Media format and control



- 6.3.1 2D video format basic set
- 6.3.2 3D format basic set
- 6.3.3 Audio format basic set
- 6.3.4 Media service discovery
- 6.3.5 Media over internet

6.3.1 2D video format basic set

It is recommended that the VMS implements the following media formats:

- profiles and levels for 2D video codecs as specified in ETSI TS 101 154 (DVB A001) [TS 101 154]

For Broadcast services over terrestrial, satellite and cellular networks it is recommended to employ 2D video formats:

- When DVB protocols are used [EN 302 755], [EN 302 307], [ETS 102 796], [TS 103 285], [DVB C100], profiles and levels for 2D video codecs are specified in ETSI TS 101 154 (DVB A001);
- When multimedia broadcast / multicast service (MBMS) protocols are used [TS 26 346], as a vehicular multimedia system (VMS), it should support the video profile operation points as specified in 3GPP TS 26.446 [TS 26.446]
- When advanced television systems committee (ATSC) protocols are used [ATSC A330], [ATSC A331], [ATSC A332], VMS should support the video profile operation points as specified in [ATSC A341].

For television services over 3GPP services, it is recommended that the VMS supports the video profile operation points as specified in 3GPP TS 26.116 [TS 26.116].

For VMS streaming services over bidirectional cellular network; it is recommended that the VMS supports 2D video formats from the following specifications:

- ETSI TS 103 285 DVB-dynamic adaptive streaming over HTTP (DASH) [TS 103 285];
- 3GPP-DASH as specified in 3GPP TS 26.247 [TS 26.247].

6.3.2 3D format basic set

For augmented reality services, it is recommended that the VMS support 3D formats as follows:

- Moving picture expert group (MPEG) video-based point cloud compression (V-PCC) codec as specified in [ISO/IEC 23009-1] can be used to convey 3D visual objects;
- The transport of V-PCC using MPEG-DASH is specified in [ISO/IEC 23090-10].

For point cloud high definition (HD)-maps and their being generated from the cloud, it is recommended that the VMS support:

- Geometry-based point cloud compression (G-PCC) codec as specified in ISO/IEC FDIS 23090-9 [ISO 23090-9]
- When the G-PCC codec is supported the transport of G-PCC as specified in ISO/IEC DIS 23090-18 [ISO 23090-18].

6.3.3 Audio format basic set

For broadcast services over terrestrial, satellite and cellular networks, it is recommended that VMS support audio codecs as follows:

- **DVB audio codecs as specified in [TS 101 154];**
- **When 3GPP MBMS protocols are supported, Enhanced aacPlus and extended AMR-WB as specified in 3GPP TS 26.446 [TS 26.446];**
- **ATSC audio codec as specified in [ATSC A342-1][ATSC A342-2][ATSC A342-3].**

For VMS streaming services over bidirectional cellular network; it is recommended that VMS support audio codecs as follows:

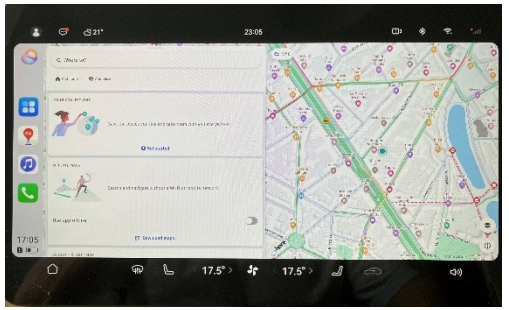
- **One or more of Enhanced aacPlus and extended AMR-WB over 3GPP packet-switched streaming service (PSS) services as specified in 3GPP TS 26.234 [TS 26.234] and over 3GPP-DASH as specified in 3GPP TS 26.247 [TS 26.247].**

6.3.4 Media service discovery

For discovering and presenting multimedia entertainment services it is recommended that VMS support DVB-I [TS 103 770]. This approach provides a user experience that is comparable with the reception of television services on any consumer electronics device.

6.3.5 Media over internet

- **Playing media from the internet requires websites to know decoding and encoding abilities of the VMS. It is recommended that the VMS support the W3C media capability standard (see: [W3C MEDIA_CAP]), which provides APIs to allow websites to get information about device decoding and encoding capabilities and to find the best match based on the device's display.**
- **The standard aims at exposing the following sets of properties:**
  - API to query a device's decoding and encoding abilities of the device based on the information (codecs, profile, resolution, bitrates) describing if the playback should be smooth and power efficient;
  - Information about display properties for selecting the right content;
  - Real time feedback about playback for allowing websites to react to CPU/GPU usage in real time.



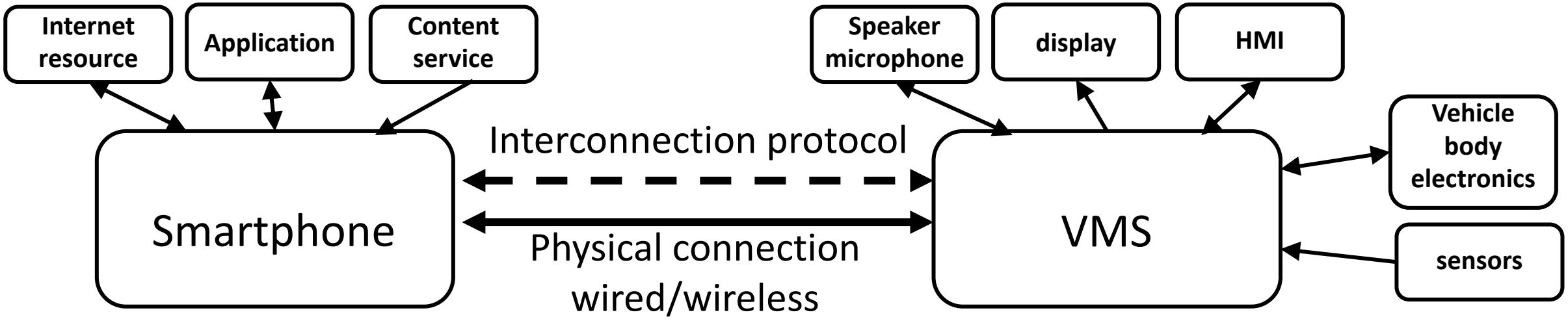
6.4 Smartphone Infotainment Assistant

6.4 Smartphone Infotainment Assistant

- 6.4.1 Introduction of Smartphone Infotainment Assistant
- 6.4.2 Connectivity requirements
- 6.4.3 Communication layers
- 6.4.4 Protocol flow diagrams
- 6.4.5 System performances

6.4.1 Introduction - architecture

Using a Smartphone Infotainment Assistant to perform Smartphone applications and functionalities with the vehicle HMI and dashboard, while sharing resources between the Smartphone and the VMS and exchanging information.
The information exchange is performed by using the interconnection protocol. It is recommended that VMS use wireless connectivity to support the interconnection protocol. It is also permitted that the VMS uses the wired connectivity alternatively or in combination with the wireless connectivity



6.4.2 Connectivity

It is recommended that the physical connection between the Smartphone and the VMS uses the following combination of short-range communication technologies:

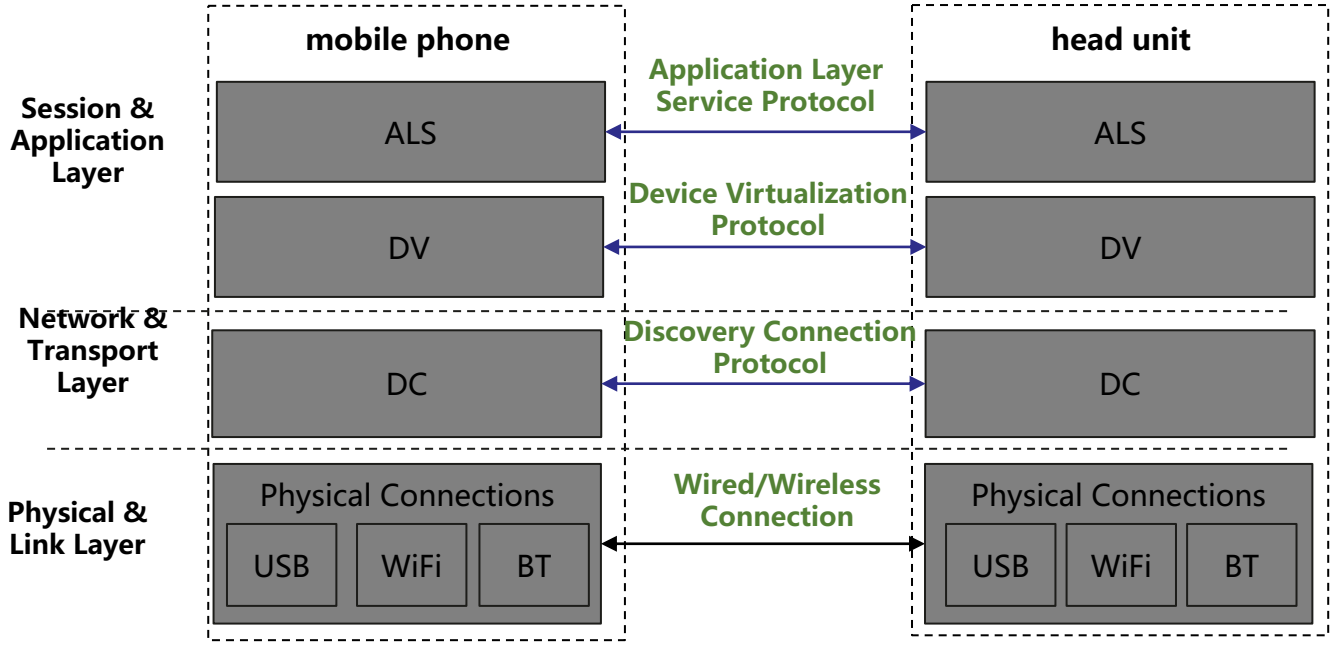
- **BT (see 6.1.2 BT (Bluetooth))**
- **WLAN (see 6.1.1 WLAN)**

It is permitted that the physical connection between Smartphone and the VMS uses the USB wired communication technology (see 6.1.4 USB)

6.4.3 Communication layers

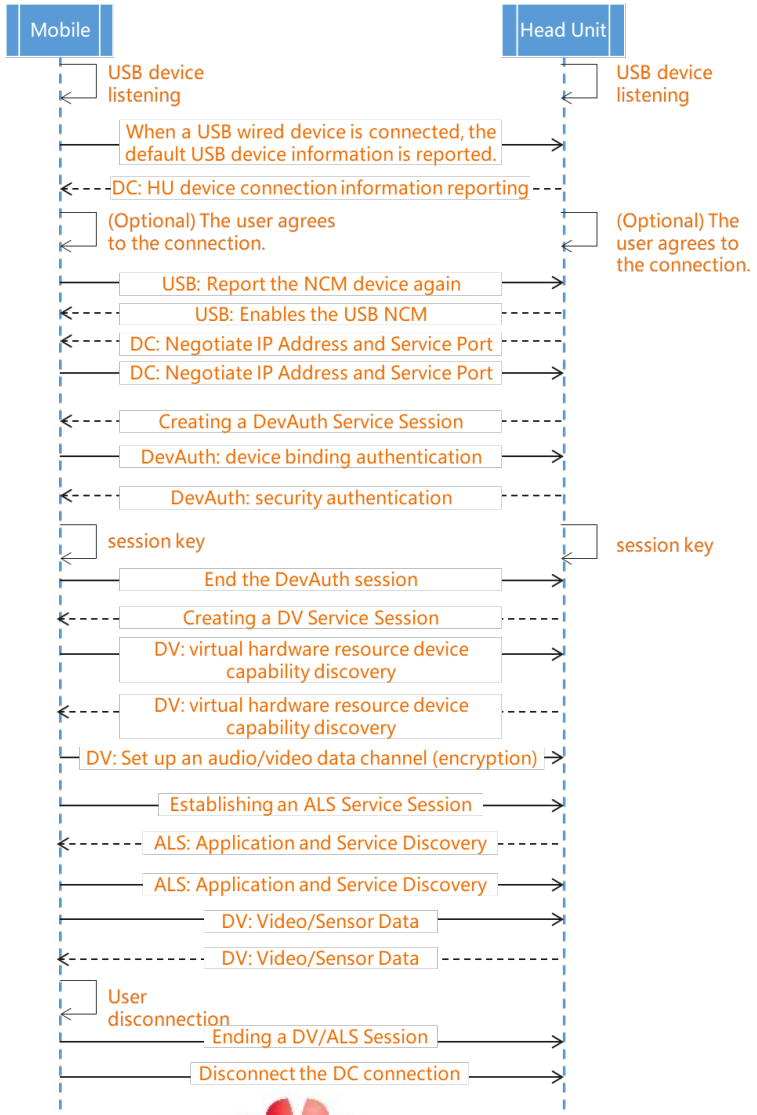
It is recommended for VMS to apply implement the following communication layers:

- **DC protocol: dealing with device discovery and connection, and establishes network transmission channels based on upper-layer service requirements.**
- **DV protocol: used to virtualize shared hardware at both ends (e.g. mobile phone sending audio data to car speakers)**
- **ALS protocol: used for service discovery, invoking, and collaboration at the application layer between devices (e.g. in-vehicle infotainment to use mobile phone internet access 4G/5G)**

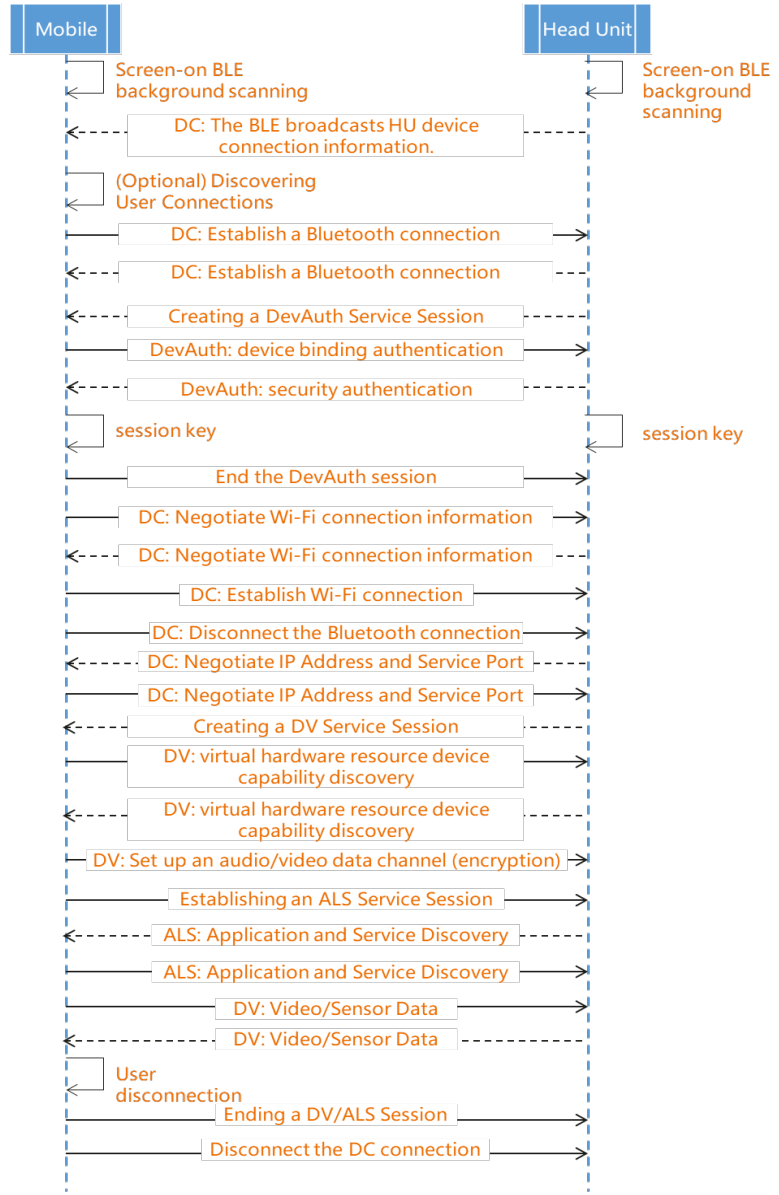


6.4.4 Protocol flow diagrams

wired



wireless



System performances

Domain	Characteristic	Criteria
Wired connection	First connection completion delay	≤ 5 s
	Non-first connection completion delay	≤ 5 s
	Connection success rate	≥ 99%
	Long-term connection reliability (disconnection times)	0 (at least 8 hours)
Wireless connection	First discovery completion delay	≤ 1.5 s
	First discovery success rate	> 70%
	Furthest discovery distance	150 cm
	First connection completion delay	≤ 5 s
	Non-first connection completion delay	≤ 5 s
	First connection success rate	≥ 99%
	Long-term connection reliability (disconnection times)	0 (at least 8 hours)
	Reconnection completion delay (which is started after vehicle shutdown)	≤ 5 s
	Reconnection success rate	≥ 99%
Projection display	Screen projection delay	≤ 150 ms
	Projection frame rate	FPS ≥ 55 (60 fps video source) FPS ≥ 24 (30 fps video source)

Domain	Characteristic	Criteria
Audio output	Audio output delay	≦1000 ms
	No frame freezing duration for audio output	≥ 4 h
	Voice and picture synchronization delay	T ∈ (−125 ms, +45 ms) ~ (−185 ms, +90 ms) NOTE – A positive value indicates that the sound is ahead of the picture, and a negative value indicates that the sound is behind the picture.
User operations	Touch screen response delay	≤ 200 ms
	Steering wheel control response delay	≤ 200 ms
Voice interaction	Voice wakeup success rate (low noise)	≥ 93%
	Voice wakeup success rate (medium noise)	≥ 90%
	Voice wakeup success rate (high noise)	≥ 85%
	Voice interaction response delay (entertainment category)	≤ 2600 ms
	Voice interaction response delay (call category)	≤ 2100 ms
	Voice interaction response delay (navigation category)	≤ 2500 ms
	Voice interaction success rate (entertainment category)	≥ 85%
	Voice interaction success rate (call category)	≥ 85%
	Voice interaction success rate (navigation category)	≥ 85%

Offline feedback

Offline feedback about voice assistant

Culturally specific features and interactions

- **We frequently see that the demand for specific features and interactions is very different between different cultures – especially between Eastern-Asia and Europe / NA. We would recommend checking the non-basic feature set provided in the voice assistant against local customs and cultural aspects.**
- **Options for customization**
  - Our participants frequently ask for the option to make the assistant more "personal" by having the option to rename the voice assistant, and to adapt its visualization to their individual taste – this is especially true for more complex assistants. The preferences for visualization are again often tied to cultural aspects. Participants also frequently ask for the option to change the voice of their assistant – most basic is from female to male – but some also ask to have a cartoon character's or a well known TV character's voice as their assistant's voice.
- **Visual feedback**
  - We often see that it helps participants to interact with the voice assistant more productively if they receive some visual feedback from the voice system that helps them with understanding what the system has understood, and what state it is in. E.g. idle, ready to receive, computing, receiving.

长安汽车 CHANGAN HUAWEI

Thank you

长安汽车
CHANGAN