WIND RIVER

Multimedia Resources in Linux for Automotive

Table of Contents

Executive Summary

Intelligent in-vehicle systems increasingly offer end users rich multimedia experiences as part of their core functionality. To compete in a dynamic marketplace, automotive original equipment manufacturers (OEMs) and integrators must provide digital audio and video, animated 2D and 3D maps and street views, and speech-based interfaces in their next-generation in-vehicle systems.

Linux, open source software (OSS), and proprietary/commercial software hosted on Linux offer automotive OEMs and suppliers a wide range of multimedia technologies, toolkits, media players, and plug-ins. This paper provides an overview of available multimedia software resources for automotive applications that include audio, video, speech, and digital rights management (DRM). When possible, it also supplies heuristics for making design and deployment decisions based on project maturity and licensing information, for example.

Core Automotive Multimedia Requirements

To support real-world automotive applications, Linux-hosted multimedia technology for automotive applications needs to deliver certain key capabilities:

- Support for popular, industry-standard media formats and file types
- Ready-to-use audio/video codecs and players
- Screen resolutions from VGA (video graphics array) (640x480) upward through high definition (HD) resolutions
- Performance and load characteristics matched to available automotive embedded hardware
- CPU support for automotive processor architectures: ARM, Hitachi SH, and Intel
- Speech output and recognition

Challenges to Meeting Multimedia Requirements with Open Source Software

Until recently, many OEMs found it extremely difficult to integrate a useful and comprehensive set of multimedia components for use in devices based on embedded Linux. Despite broad embedded and desktop deployment and a worldwide community of Linux developers and end users, efforts to accommodate the gamut of media types, codecs, formats, DRM regimes, intellectual property restrictions, and licenses stymied the deployment of full-featured Linux-based media devices. These challenges include the following:

- Highly proprietary, nonstandard nature of many audio and video codecs
- Variation in maturity of OSS projects implementing multimedia components
- Conformance to the particulars of copyright legislation (in particular the Digital Millennium Copyright Act, or DMCA) and patent law in choosing and integrating OSS components for multimedia

This paper, in surveying available multimedia technologies for embedded Linux, refers to relevant information regarding maturity, licensing, and intellectual property status of the software in question.

Use Cases and Applications

The primary use cases for multimedia technology in Linux-based automotive applications include the following:

- Vehicle dashboard and seatback displays with medium to high resolution
- Navigation systems with 2D and 3D map displays, audio output, text-to-speech, etc.
- Entertainment systems for local media playback and streaming
- Operational feedback (e.g., vehicle reverse video camera)

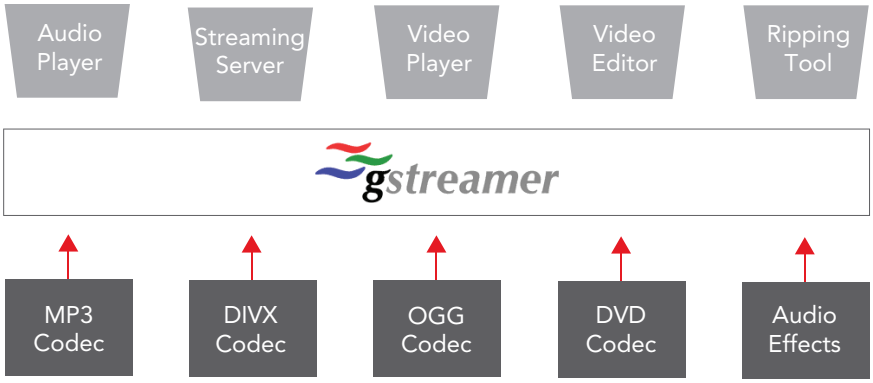

Figure 1: GStreamer plug-in and application architecture

Multimedia Infrastructure: Frameworks, Engines, Etc.

As with many other Linux-based technologies, there exist multiple, sometimes competing OSS projects to implement multimedia functionality as well as versions of cross-platform commercial technology targeted and optimized for Linux-hosted execution.

ALSA

ALSA, the Advanced Linux Sound Architecture, provides standard audio and MIDI services to the Linux operating system and applications running on it. ALSA supports a range of audio interfaces, from PC sound cards to professional-quality multichannel audio. ALSA offers features advantageous to in-vehicle infotainment use:

- Dynamically loadable, modular audio device drivers
- Native support for symmetric multiprocessing (SMP) and thread safety
- alsa-lib library, with user-space APIs to simplify application programming
- Support for legacy open sound system APIs, with binary compatibility for most Open Sound System programs

ALSA is part of the Linux kernel and is covered by GPLv2 and the Linux COPYING file. alsa-lib is distributed under the GNU Lesser General Public License (LGPL), easing OEM integration. Learn more at http://www.alsa-project.org.

GStreamer

GStreamer is a multimedia framework serving a range of applications, including media players, streaming media broadcasters, and video editors. GStreamer is widely deployed in both desktop and embedded settings and has been adopted by mobile and embedded Linux platforms, including Maemo, Moblin, and LiMo.

GStreamer connects media processing elements into a pipeline, where each element is provided by a plug-in. Elements communicate by means of pads (like solder points on silicon). A source pad on one element can be connected to a sink pad on another. When the pipeline is in the playing state, data buffers flow from the source pad to the sink pad. Pads negotiate the kind of data that will be sent using abstractions for defining capabilities.

GStreamer lets developers integrate codecs and filters through plug-ins with generic interfaces. By decoupling codecs from the framework architecture, multimedia applications can take advantage of advances in codec and filter technology asynchronously from GStreamer development.

GStreamer is released under the LGPL, facilitating integration and commercial deployment with minimal reciprocal disclosure. Learn more about GStreamer at http://gstreamer.freedesktop.org/.

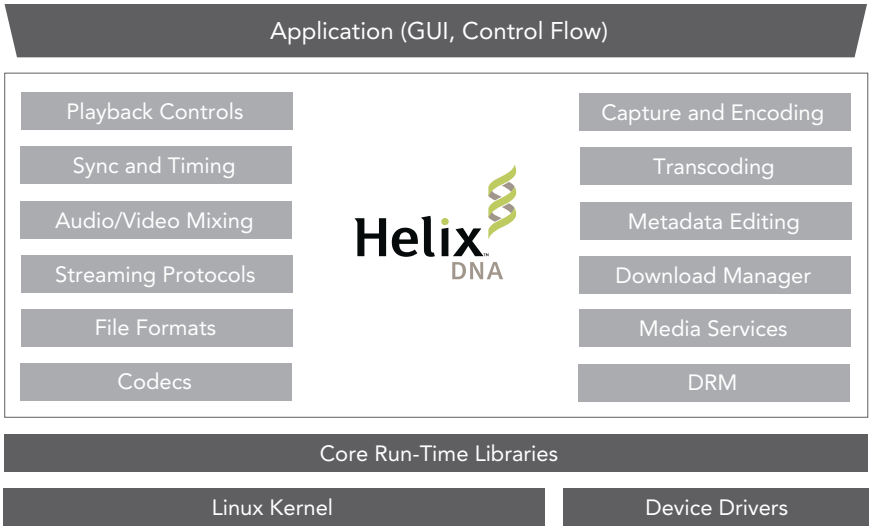

Figure 2: Helix DNA Client system architecture

Helix DNA Client

Helix DNA Client is a "universal" open source digital media playback engine, supporting playback of any format and codec across a range of device types. Helix DNA provides underlying multimedia functionality to a range of applications requiring local playback and streaming of multiple formats in disparate environments.

Helix DNA Client underlies the digital media platform for RealNetworks' RealPlayer on Windows, Apple, and Linux as well as the Rhapsody player.

Helix DNA and codec source code is available for reference and research and development under Open Source Initiative (OSI)-approved licenses (GPLv2 and RealNetworks Public Source License) or via commercial licenses for deployment in devices. Learn more about Helix DNA Client at https://helix-client.helixcommunity.org/.

Adobe Flash Lite

Adobe Flash Lite software provides a run-time engine for mobile and consumer electronics devices with substantial interoperability benefits for device manufacturers and content developers. Automotive OEMs and integrators can build on the popularity and ubiquity of Adobe Flash content to provide end users with web browsing, video, and interactive content comparable to users' desktop experiences, in the car, on the go.

Adobe Flash Lite offers the following to Linux-based automotive systems developers:

- FLV file type support, including H.264, On2 VP6, and Sorenson video codecs
- Access to content and video created with Adobe Flash
- Integrated authoring environment
- Cross/multiplatform support

Flash Lite features an open SDK but is licensed under strictly commercial terms by Adobe for device-based deployment. To learn more, visit http://www.adobe.com/products/flashlite/ for general information and http://www.adobe.com/products/flashlite/version/ for a comparison of Adobe Flash products, capabilities, and platform support.

Xine

Xine is a multimedia playback engine built around the shared library xine-lib. Xine offers a variety of front-end player applications, including Codeine, gzine, Kaffeine, KMPlayer, and Totem.

Xine can open and execute container formats that include 3gp, AVI, ASF, FLV, Matroska, MOV (QuickTime), MP4, NUT, Ogg, OGM, and RealMedia. Xine can decode/play a range of audio and video formats:

- **Audio:** AAC, AC3, ALAC, AMR, FLAC, MP3, RealAudio, Shorten, Speex, Vorbis, WMA
- **Video:** Cinepak, DV, H.263, H.264/MPEG-4 AVC, HuffYUV, Indeo, MJPEG, MPEG-1, MPEG-2, MPEG-4 ASP, RealVideo, Sorenson, Theora, WMV (partial, including WMV1, WMV2, and WMV3, via FFmpeg)

Xine is released under the GNU General Public License and may not be suitable for all commercial applications. Learn more about Xine at http://xinehq.de/.

Open Source Media Formats

Current media formats on Windows desktop machines and ubiquitous media devices (MPEG, WMV, etc.) are probably familiar. But the range of freely licensed media formats supported by Linux-based frameworks and players is probably less familiar. These formats provide additional options for OEMs to offer high-quality multimedia experiences without the obligation of patents, proprietary licenses, and so on. However, they are not as widely used as their proprietary equivalents. Note that while these and other open formats are native to open source multimedia software, OSS also supports familiar commercial formats through a mix of open source and proprietary plug-ins and codecs.

Ogg Vorbis

Ogg Vorbis is an open, patent-free, audio-encoding and compression format that also offers streaming technology. It is comparable to other formats used to store and play digital music, such as MP3, VQF, AAC, and so on. Ogg Vorbis differs from other formats by being completely free, open, and unpatented. Ogg is the native audio file format for many open source media tools and supported by an increasing number of commercially available media players and devices as well.

Ogg Vorbis comprises two components:

- **Ogg:** Container format for audio, video, and metadata from Xiph.org
- **Vorbis:** The audio compression scheme contained in Ogg; other formats can be embedded in Ogg such as FLAC and Speex

Learn more at http://www.vorbis.com.

Theora

Theora is a free and open video compression format. Theora scales from postage stamp to high-definition resolution and is competitive with commercial formats, especially at low bit rates. Theora is supported by most open source engines and players, including MPlayer, Xine, Helix, and VideoLAN. Theora is comparable to MPEG-4 and DiVX, with a large and active developer community engaged in incremental enhancement.

Theora can be used to distribute film and video online and on disc without licensing or royalty fees or vendor lock-in associated with other formats. Learn more at http://www.theora.org and http://www/xiph.org.

Media Players

Gnash

Gnash is a free software Flash movie player and is one of the GNU family projects hosted and supported by the Free Software Foundation. Gnash is based on GameSWF and supports most SWF v7 features and some SWF v8 and v9 features.

Gnash features include the following:

- **Standalone and plug-in operation:** Gnash can run in either mode.
- **Browser-based execution:** Gnash can run as a plug-in from within Mozilla-derived browsers, such as Firefox, as well as Konqueror.
- **SWF v7+ compliant:** Gnash can run SWF-compatible files.
- **Streaming video support:** Gnash supports viewing of streaming video from popular sharing sites such as Lulu.tv or YouTube.
- **XML message server:** Gnash supports XML-based messaging for the Flash format.
- **High-quality output:** Gnash uses OpenGL for rendering graphics on the desktop, and Antigrain Geometry (AGG) for embedded frame buffer–only devices.
- **Extensible:** Gnash supports extending ActionScript with custom scripts or wrappers for third-party libraries.

As part of the main GNU project tree, Gnash is licensed under GPLv3. Learn more about Gnash at http://www.gnu.org/software/gnash/.

Helix Player

Helix Player is an open source media player widely deployed on both desktop Linux and embedded Linux, Solaris, and Symbian OS. Helix Player is based on the Helix DNA Client media engine. Helix Player is dual-licensed under open source (RealNetworks Public Source License) and commercial community source (RealNetworks Community Source License) licenses. Helix Player contains only open source components and plays open source formats, such as Ogg Vorbis and Theora.

RealPlayer for Linux builds on top of the Helix Player and supports RealAudio, RealVideo, Windows Media, MP3, Flash, MPEG-4 (fee-based), and other non–open source components.

Learn more about both Helix Player and RealPlayer at https://player.helixcommunity.org/.

Kaffeine

Kaffeine is an open source, full-featured media player for KDE (K Desktop Environment) and accompanies KDE in deployment in enterprise workstation and home desktop settings. It supports a broad range of local and network media and digital video broadcasting (DVB). Kaffeine (like Totem below) can use xine-lib or GStreamer as a back end.

Kaffeine is licensed under GPLv2 but as an application imposes minimal reciprocal disclosure requirements upon OEMs. Learn more about the Kaffeine player at http://kaffeine.kde.org/.

Totem

Totem is the standard movie player for the GNOME desktop environment and widely available with desktop Linux distributions. Totem is tightly integrated with GNOME and its file manager, Nautilus, featuring the generation of video file thumbnails and a video plug-in for Mozilla browsers (Firefox, et al.).

Totem exists in two versions: one based on GStreamer that features greater extensibility and supports a larger variety of media format, the other based on Xine that offers better encrypted DVD playback support and slightly broader file format support than the GStreamer version.

Building on GStreamer, Totem can play all mainstream media formats, both open and proprietary. It also understands numerous playlist formats, including SHOUTcast, M3U, XML Shareable Playlist Format (XSPF), SMIL, Windows Media Player, and RealAudio. Totem supports full-screen video playback including multihead configurations.

Totem is licensed under GPLv2 but as an application imposes minimal reciprocal disclosure requirements upon OEMs. Learn more at http://www.gnome.org/projects/totem/.

LinDVD

LinDVD from Corel is a software DVD player from the same team that created the ubiquitous WinDVD for the Windows desktop. LinDVD is packaged as a custom product/services offering targeted at device OEMs and is not a COTS (commercial off-the-shelf) product per se.

MPlayer

MPlayer is a movie player structured as a command-line application with GUIs for each supported operating system. For Linux, MPlayer offers GNOME-based gmplayer.

MPlayer supports a wide range of formats and encoding standards, including MPEG/VOB, AVI, Ogg/OGM, VIVO, ASF/WMA/WMV, QT/MOV/MP4, RealMedia, Matroska, NUT, NuppelVideo, FLI, YUV4MPEG, FILM, RoQ, and PVA files. MPlayer lets end users play VideoCD, SVCD, DVD, 3ivx, DivX 3/4/5, WMV, and H.264 movies.

MPlayer supports a gamut of output drivers, including X11, Xv, DGA, OpenGL, SVGAlib, fbdev, AAlib, DirectFB, GGI, SDL, VESA, and various low-level card-specific drivers (e.g., Matrox, 3Dfx, and ATI). In many cases, MPlayer also supports software and/or hardware scaling.

Most video and audio formats are supported in MPlayer natively through the libavcodec library of the FFmpeg project. For formats lacking open source codecs, MPlayer relies on binary codecs. It can use Windows DLLs directly with the help of a DLL loader.

MPlayer is licensed under GPLv2 and, as an application, imposes minimal reciprocal disclosure requirements upon OEMs. However, the informal use and availability of (reverse-engineered) MPlayer codecs for otherwise proprietary media formats may expose OEMs or suppliers to intellectual property claims against use of software patents in open source software. Learn more at http://www.mplayerhq.hu/.

VLC and VideoLAN

VLC is a free software media player developed under the VideoLAN project. Monolithic in size, it encompasses a portable multimedia player, encoder, and streamer in a single code base (the front-end and plug-in approaches of GStreamer, Xine, etc.). VLC can stream content over networks and transcode media into various different formats, with support for many audio and video codecs and file formats as well as DVDs, VCDs, and streaming protocols.

VLC includes a large number of free codec libraries, reducing the need for integration and tuning of proprietary plug-ins. Many VLC codecs ship with libavcodec from the FFmpeg project. VLC is also the first player to support playback of encrypted DVDs on Linux with the libdvdcss DVD decryption library.

VLC is licensed under GPLv2 but as an application imposes minimal reciprocal disclosure requirements upon OEMs. Learn more at http://www.videolan.org/.

Speech Technology

The commercial marketplace, academia, and the open source bazaar are flush with speech technologies. While many are quite mature and stable, few have the quality and performance to make the cut as "automotive grade."

Most of the projects and products available for Linux for automotive focus individually on speech output (speech synthesis and text-to-speech) or input (recognition).

Speech output for in-car systems can be a key differentiator in a crowded marketplace. Requirements include smooth, continuous speech generation, choice of voices, and off-the-shelf internationalization capability.

Festival

Festival offers a framework for building speech synthesis systems and offers full text-to-speech (TTS) through multiple APIs and channels: shell/CLI, Scheme command interpreter, C++ and Java class libraries, and Emacs interface. Festival supports English (British and American) and Spanish, with English being the most mature. Tools and documentation for new voices are available through Carnegie Mellon University's Festvox project (http://festvox.org).

Festival and the speech tools are distributed under an X11-type license, allowing unrestricted commercial and noncommercial use alike. Learn more at http://www.cstr.ed.ac.uk/projects/festival/.

MBROLA

The MBROLA project creates speech synthesizers for a growing list of languages and is provided free for noncommercial applications. The project goal is to foster research into speech

synthesis, specifically on prosody. As such, MBROLA is not a TTS synthesizer; rather it accepts phoneme representations as input, together with prosodic information, and produces speech output based on concatenation of diphones.

MBROLA is distributed under its own licensing terms (essentially a limited free copyright license). Learn more at http://tcts.fpms.ac.be/synthesis/mbrola.html.

Nuance RealSpeak Solo

RealSpeak Solo is a TTS solution, optimized to enhance automotive and other embedded conversational applications. Highly scalable, it provides high speech quality output across a range of footprints, typically from 8MB to 20MB for automotive use. Recently, Nuance introduced the option for voice capture of preferred speakers for greater OEM differentiation.

Nuance offers several additional products for in-vehicle speech generation, including PromptSculptor and SpeechPAK Automotive. Learn more about Nuance products at http://www.nuance.com/automotive/.

Speex

Speex is an open source, patent-free audio compression format designed for speech. The Speex project aims to lower the barrier of entry for voice applications by providing a free alternative to expensive proprietary speech codecs. Moreover, Speex is based on CELP (Code Excited Linear Prediction) and can compress voice at bit rates from 2kbps to 44kbps. Some of Speex's features include the following:

- Narrowband (8kHz), wideband (16kHz), and ultra-wideband (32kHz) compression in the same bit stream
- Intensity stereo encoding
- Packet loss concealment
- Variable bit rate operation (VBR)
- Voice activity detection (VAD)
- Discontinuous transmission (DTX)
- Acoustic echo cancellation and noise suppression

Speex is part of the GNU Project and is available under the revised Berkeley Software Distribution (BSD) license. Learn more at http://www.speex.org/.

NeoSpeech VoiceText

VoiceText generates synthesized speech from input text and targets custom, standalone applications. VoiceText is highly scalable, with footprints as small as 16MB. VoiceText supports text-to-speech for U.S. English, Latin-American Spanish, Korean, Japanese, and Mandarin Chinese, with a collection of 11 native voices.

NeoSpeech supports desktop Linux off-the-shelf and embedded Linux in custom configurations. Learn more at http://www.neospeech.com.

Voice Recognition

Vehicle-based speech recognition is one of the more exacting uses of this input technology: viable recognition systems require high-quality noise rejection and compensation to offset high levels of ambient noise (wind, traffic, street sounds). Voice recognition systems must be speaker-independent and offer low recognition latency and high accuracy. As such, only a few of the many speech input technologies available as open source and/or Linux-based commercial software can pass muster as automotive grade.

Nuance VoCon

VoCon 3200 is designed for adding speech recognition to automotive and other embedded applications. The VoCon 3200 speech recognition engine delivers speaker-independent and continuous speech-recognition capabilities and is suitable for navigation systems, telematics applications, and voice control of in-vehicle devices.

Modular and scalable, VoCon 3200 meets the requirements of different applications and markets with modules for speaker adaptation, natural language understanding, spelling, and online grammar compiler. VoCon 3200 features include the following:

- Speaker-independent, continuous-digit, and command recognition
- Voice destination entry for large countries and markets
- Automotive acoustical models and automatic noise cancellation

The VoCon 3200 package includes support for U.S. English, U.K. English, and German with the automotive acoustical models. Other languages can be supported easily via data files available from Nuance.

Learn more about VoCon 3200 at http://www.nuance.com/vocon/3200/.

Sensory FluentSoft

Sensory's FluentSoft technology is capable of recognizing thousands of words. Designed to create a natural human-machine interface (HMI), FluentSoft employs word and phrase spotting technologies to avoid the limitations of discrete word command and control.

To aid in speech application development, Sensory offers a speaker-independent software development kit (SDK) with a large vocabulary for integrating high-quality speech recognition technology into automotive applications, with modest memory and CPU resource use. The FluentSoft phonetic recognition engine can build new vocabularies based on text input, even on the device.

Sphinx

Sphinx is an open source speech recognition system implemented in Java. Sphinx comes from a collaboration among Carnegie Mellon University, Sun Microsystems Laboratories, Mitsubishi Electric Research Labs (MERL), and Hewlett-Packard (HP), with contributions from the University of California at Santa Cruz (UCSC) and the Massachusetts Institute of Technology (MIT).

Sphinx features live and batch mode speech recognizers capable of recognizing discrete and continuous speech. It also offers a range of plug-in interfaces for different front ends, language models, and acoustic models. Sphinx also boasts standalone tools for displaying waveforms and spectrograms and generating features from audio.

Sphinx is distributed under a BSD license. Learn more at http://cmusphinx.sourceforge.net/.

Digital Rights Management

DRM software and related technologies that enforce licensing do not sit easily with open source software and developer communities that create and support OSS. The latest version of the GNU General Public License (GPLv3) even features explicit prohibition against placing GPLv3 software itself under the control of a rights management regime. However, deployment and use of DRM, while incompatible with managing GPL software itself, is viable for other software and for multimedia content. For their part, OEMs, content owners, and network operators express real requirements for DRM, and a number of solutions are available for Linux-hosted automotive applications.

Helix DRM

Helix DRM for Devices enables chipset and device manufacturers to include DRM in intelligent device software. Helix DRM for Devices allows content rights holders to determine device delivery policy while also serving the needs of device manufacturers and end users. Helix DRM for Devices gives content rights holders and device manufacturers flexibility over their business models through two modes of operation.

Primary Device Support

A primary device is a network-connected system, as with PCs, set-top boxes, home media servers, or mobile phones. A primary device can connect directly to the Internet and features local storage. With Helix DRM primary device support, both the Helix DRM software and the Helix DNA client run natively on the managed device. To access and play managed content, an end user or program must acquire the secured content from a media server. When the attempt is made to use the encrypted file on the primary device, Helix DRM requires a license key to decrypt it. In response, the Helix DRM client on the primary device requests a license, and the Helix DRM license server generates a license key per the associated business rules and policy defined by the content owner.

Secondary Device Support

A secondary device requires a connection to a primary device for the transfer of data and rules for that data. On secondary devices, Helix DRM lets OEMs enable standards-based secure subscription, online purchase, and other business models on a wide range of portable devices and connected home appliances. Learn more at https://devicedrm.helixcommunity.org/.

OpenIPMP

OpenIPMP is an open source DRM implementation for MPEG-4 and MPEG-2 content. The project adheres to ISO/MPEG IPMP (intellectual property management and protection) open standards (MPEG IPMP hooks and IPMP-X), ISMAcryp, and Open Mobile Alliance (OMA) DRM 2 specifications. OpenIPMP includes CA, PKI, DOI management, ISMA streaming, license server, encoding/encryption, player, and plug-in architecture.

The OpenIPMP project code is distributed under the Mozilla Public License (MPL 1.1) and so easily integrates into commercial/proprietary device software stacks. Learn more at http://sourceforge.net/projects/openipmp.

Conclusion

Thanks to substantial investment by the automotive and consumer electronics industries, academic researchers, and open source community work and through the efforts of commercial independent software vendors (ISVs), embedded Linux today hosts a rich and varied toolbox of multimedia, speech, and DRM technologies. This document is not intended to be exhaustive but rather to provide an overview of the options in these areas that are appropriate for automotive development and deployment. In many cases, the software technologies and commercial offerings highlighted here are already integrated by Wind River or deployed by global automotive OEMs.

Notes

1. GNU General Public License version 3, Section 3, http://www.fsf.org/licensing/licenses/gpl.html.