January 27, 2012

Whitepaper

IP VIDEO DOOR PHONE

Contents

P.no

INTRODUCTION

Audio Video Door Phone Systems are fast becoming de rigueur in office buildings, residential complexes, villas and bungalows. Technological advances in multimedia and embedded systems in general such as camera sensor technology and real time video encoding & decoding technologies, have made it commercially viable for current day security solutions to incorporate video and become a true multimedia solution.

Calsoft Labs is a product engineering services company that specializes in the three niche areas of Networking and Datacom, Embedded Systems and ISVs. Calsoft Labs is uniquely positioned to offer its customers the best of both worlds – the personalized services offered by a small company combined with the technical know-how and capability of a larger company.

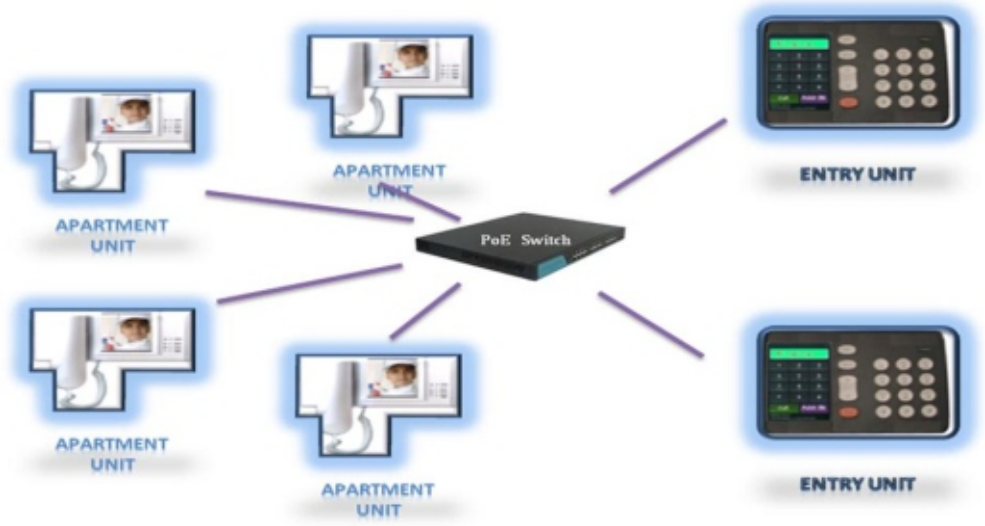
As a one stop solution house for embedded products, Calsoft Labs' typical engagement ranges from product requirement scoping, system design and development to deployment and maintenance including the mechanical and plastic components of the solution. On the technology front, we have expertise in a wide range of domains from high-speed hardware boards design, to firmware, OS, driver and other components of board support packages. This also includes design and development of target applications and host side applications that generally run on PCs.

IP Video Door Phone System Overview

The IP Video Door Phone System supports video capture and streaming from Entry Units to Apartment Units as well as two way audio conversations via an on-board microphone and speaker. One can make a call to an Apartment Block via Entry Units placed at the doors and the call can be answered by any of the Apartment Units in that block.

Both Entry Units and Apartment Units have LCD displays with built-in Touch-Screen interface and keypads. An Entry Unit is mounted with a CMOS camera module and the SoC of both units have integrated hardware accelerators for video encoding and decoding. Thus one can experience real time video streaming at full frame rate from Entry Units to Apartment Units. The LCD (customizable to required resolution) and the touch screen interface provide a rich GUI experience similar to those in smart phones available today. The presence of an AEC (Acoustic Echo Cancellation) chip on the audio path handles any acoustic echoes that are generated by such a system and ensures that the listener will never hear the echo of his own voice. The Apartment Units may support a standard off-the-shelf telephone handset, or can be hands-free.

The GUI subsystem is based on QT 4.6 which is an open-source distribution under Nokia support or GTK. These are proven platforms that provide rich set of application building blocks and deliver all of the functionality needed to build advanced cross platform applications. The IP Video Door Phone System is a scalable system and supports various deployment scenarios - such as one unit to multiple units or one to one or many to many units.



System Block Diagram

Both Entry Units and Apartment Units are based on the same ARM based, multimedia System-on-Chip (SoC), with hardware acceleration for image processing, video encode and decoding. SoC also includes most of the peripheral controllers built within. Both units use DDRAM based memory and NAND Flash based storage.

The Apartment Unit comprises of a 240x320 LCD with touch screen, fast Ethernet, microphone, ringer circuit / handset, keypad, LEDs and PoE features. Optionally it can include a speaker, camera module and composite video interface. It can be customized to be powered by external power adapters instead of PoE.

The Entry Unit comprises of an LCD with touch screen, fast Ethernet, microphone, speaker, camera, keypad and PoE features. It can also comprise interfaces with additional sensors and control devices for extended functionality.

Features and Functionalities ▶▶▶

The IP Video Door Phone system is a feature-rich solution that matches the best in the market. These features have been developed after extensive consultation with domain experts who constantly interact with designers, manufacturers, resellers and end user.

Features and functionalities of Entry Unit:

A. Video:

- Video capture - portrait mode, 240x 320 resolution, optionally scalable to higher resolution
- One way peer-to-peer full frame rate video streaming using RTP/RTSP protocol
- Support for MPEG4 (or H.264) standard video compression algorithm
- Image capture and storage in JPEG format along with Call Log
- Audio Video synchronization

B. Audio:

- G.711 audio encoding/decoding
- Two way audio streaming
- Hands free audio operation
- Support for hardware based Acoustic Echo Cancellation (AEC) ensuring echo-free full duplex conversation

C. GUI:

- QT based GUI based UI rich in look and feel
- Portrait mode 240 x 320 resolution TFT LCD display, optionally scalable to higher resolution
- Resistive touch screen interface OR Keypad input, selectable at installation time
- Intuitive GUI - Dial Pad, Call Progress window, Address Book with list of Apartment Blocks with name and number

D. Call Manager:

- Ability to initiate calls to any Apartment Block
- Ability to ring and stream video to any unit
- Initiate call using Apartment number
- Initiate call using Apartment name
- Menu to choose apartment name/number to call

E.Peripherals:

- Push button Key-Pad for dialing
- Back light control for the key pad
- LED based light to illuminate object / person during night
- Touch-screen (resistive control) based key functions
- Back light control for LCD display – LCD display is ON during use and is OFF when not in use.
- Interface to several sensors and control devices for additional features

Features and functionalities of Apartment Unit:

A. Video:

- Playback of full frame rate (@30 FPS) video stream from Entry Unit
- Video Playback - portrait mode, 240x 320 resolution video display on LCD, optionally scalable to higher resolution
- Optional support for composite video display supporting NTSC/PAL, scalable to 720P
- Support for MPEG4 (or H.264) standard video compression algorithm
- Audio Video synchronization

B. Audio:

- G.711 audio encoding/ decoding
- Two way audio streaming
- Support for off-the-shelf standard telephone handset to ensure privacy of apartment residents

C. GUI:

- QT based GUI rich in look and feel
- Portrait mode 240 x 320 resolution TFT LCD display, optionally scalable to higher resolution
- Optional support for composite video display supporting NTSC/PAL, scalable to 720P
- Display of missed caller image along with call log
- Intuitive GUI menu to navigate call history details and associated snapshot images
- Host Windows PC Tool capable of downloading image snapshots and call logs

D. Call Manager:

- Ability to receive calls from any Entry Unit
- Incoming calls ring and stream video on all the idle Apartment Units within the Apartment Block
- System can support feature rich call switching functionality
- Support for two-way audio calls between Apartment Blocks

E.Peripheral Functions:

- Push button functions or touch screen for user inputs
- Ability to capture images
- Ring tone generation

Optional interfaces with third party devices for enhanced functionality and integration Common Functionalities of AU/EU:

A. Diagnostic Functions:

- Log file with extensive diagnostic information
- Simple boot time diagnostic functions with log messages on a RS232 debug port
- Ability to log all the activities for user functions as well as audit trail Support for downloading log file to Windows host using the PC Configuration/Management Tool

B. Network Functions:

- Transport stream based on widely used industry standard RTP /RTSP. Ensures interoperability with third party access control system using same transport mechanism
- Time synchronization of the system using NTP –

Windows PC Based Configuration Tool

- Host Windows based Configuration Tool for Configuration installation and administrative purposes
- Support for firmware upgrade over network using Configuration Tool
- Assignment of IP addresses to AU/EU over network using Configuration Tool
- Downloading to Windows PC, of call AU/EU logs, snapshot images over network using Configuration Tool

Design Considerations

Demands of Video Processing

Video processing involves complex and computationally intensive Video Encode/ Decode compression algorithms that need to be executed in real time under demanding time constraints. This is apart from host of other algorithms required to carry out image processing functionalities on captured image data from camera sensor to improve the quality of the image. In addition to these inherent complexities, the tasks have to be carried out in real-time. For better viewing experience, video needs to be streamed and played back at higher Frames per Second (FPS).

Video Encode / Decode – Clock requirement

Embedded application processors, like ARM, are designed to cater to a wide range of general embedded applications. Most of these applications are not computationally intensive mathematical algorithms like video encode and decode. However as a general application processor, it does support video encode and decode. It is sometimes used for such applications, but usually at the lowest resolutions and FPS, though at a cost that results in a greater load on the CPU.

A study by a group on ARM architecture indicates that the clock requirement on ARM920T for MPEG-4 algorithms for QCIF (176 x 144 pixels - Quarter CIF) at 15 FPS is 47 MHz for encode and 10 MHz for decode. One should note that these algorithms do not cover many other algorithms involved in image processing.

QCIF is meant for mobile handsets and cannot be used for Video Door Phone applications as the latter requires higher resolutions and higher FPS. Extrapolating the above clock measurements for QVGA (320x240) at 30 FPS one will get numbers like 280 MHz and 90 MHz. This clearly rules out the possibility of running the video encoder and decoders on ARM and indicates the need for dedicated hardware accelerators to perform video encode and decode compression algorithm operations. Further image processing algorithms will also need such modules in the hardware. These hardware accelerators need to be tightly integrated with the other video, DMA, memory subsystem within the CPU. This makes the choice of CPU critical for the IP Video application.

Demands of Embedded System on CPU

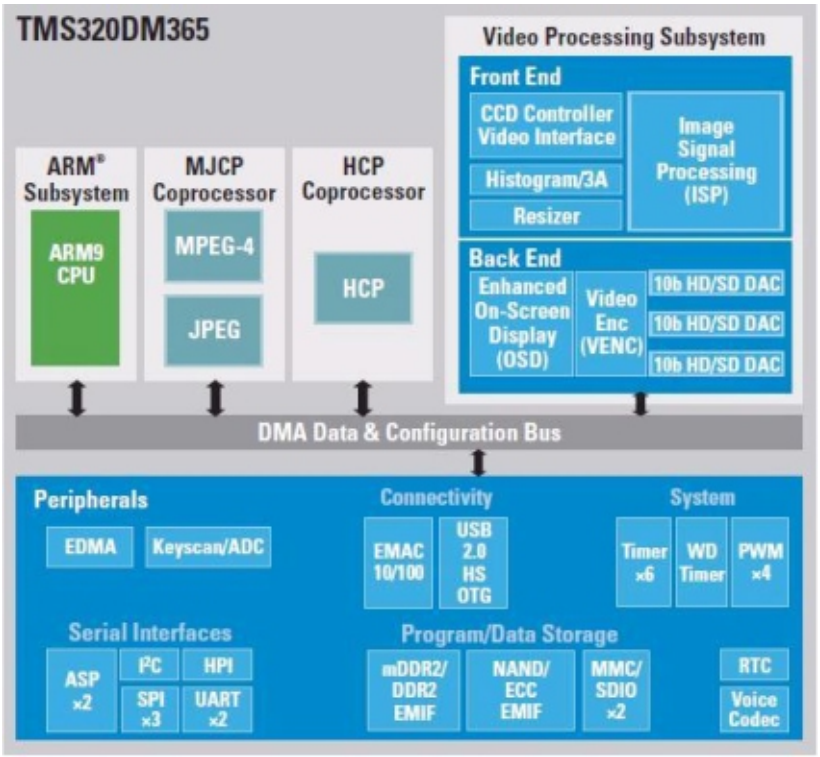
Embedded applications typically require interactions with other systems through various interfaces such as Fast Ethernet, RS232, USB etc. They need to support non-volatile storage apart from volatile memory required for system execution. One should be able to control various relays and also be able to sense various environmental parameters like ambient light, presence of persons etc. through the use of sensors. This requires that it be possible to interface relays and sensor easily. Peripheral interfaces for display and key input are a must. For timekeeping it will have to support RTC. It should also be possible to interface with audio components like microphone and speaker easily.

A reliable system cannot assume that a system is completely free from malfunctions. However minute and remote is possibility, a reliable system should include a mechanism by which the system automatically comes back to a known state, possibly by way of restarting itself. For this, it should include Watch-Dog-Timer (WDT) controller.

For obvious reasons of simplicity, cost/size reduction and reliability, these features should be inherently supported by the SoC - CPU chipset, through various built-in peripheral controllers.

TI DM365 – Digital Multimedia SoC

IP Video Door Phone System is based on TI's DM365 SoC. It is a Digital Multimedia System on Chip whose CPU core is ARM926EJ-S. Apart from including interfaces for all the peripherals required for building an embedded system such as memory, Ethernet, USB, UART, LCD, keypad, GPIO, WDT, RTC, Voice Codec etc, it also includes hardware accelerators for video encode-decode and image processing algorithms.



Salient Features of DM365:

- ARM926EJ-S™ Core at 216, 270, 300 MHz
- 10/100 Mb/s Ethernet Media Access Controller EMAC
- Memory interface: DDRAM, NAND Flash
- Key Scan, USB 2.0 Host / OTG, SD/MMC, GPIO, I2C, SPI
- Voice Codec
- Video Co-Processor
- Video Processing Front End
- Video Processing Back End

Video Co-Processor:

It includes two Video/ Image Co-Processor hardware accelerators – HDVICP and MJCP, which together support a range of video / image encode and decode operations. The algorithms supported are H.264, MPEG4, MPEG2, MJPEG, JPEG, WMV9/ VC. The resolution supported scales from QCIF to NTSC/PAL, 720P and 1080i. The maximum FPS supported depends on the resolution, though for midsize resolutions, full frame rate should not be a problem.

Video Processing Front-End (VPFE):

The VPFE port provides an interface for CMOS imager camera modules. It also includes various sub modules required for image processing that improve the quality of video image.

Video Processing Back-End (VPBE):

The VPBE provides hardware On Screen Display (OSD) support and interface for both digital LCD and analog display units through composite video interface. The analog interface supports both SDTV (NTSC/PAL) and HDTV displays.

Acoustic Echo Cancellation

Hands-free operation of communication devices is a feature that is always good to have. Since the door Entry Units are operated by general public, making it hands-free minimizes the possibility of system breakdown due to careless operation by the user.

However having a hands-free audio system poses a unique challenge. In such systems, whatever is played on the speaker (i.e. voice of the far-end person) is picked-up by microphone and fed to the audio input. There is no way for the microphone to distinguish the sound from the speaker at the other end with sound made by person standing in-front of it. This signal, unlike noise, has the same characteristics as the regular signal and hence is streamed back to the other end (i.e. AU) as an annoying echo.

This acoustic echo can be heard by the speaker at the other end if the round-trip delay is high enough for the human ear to perceive it as a distinct sound. This delay is in the order of few hundred milliseconds. In case of long distance calls, this delay is inherent to the transmission channel and is known as transmission delay.

In the case of the Video Door System where both Entry Unit and Apartment Unit are physically connected to the same LAN, this transmission delay is in the order of few tens of milliseconds and will not cause the returned sound to be perceived as an echo. However, a more complex system with audio-video streaming and playback which also supports several other functionalities including GUI, demands a multi-tasking OS as its underlying platform to carry out all the activities. Latency is inherent in such a system. This is nothing but the delay observed before the task is actually performed completely.

If we consider all the tasks that get involved in a round-trip of voice signal from AU-to-EU-back-to-AU, this latency delay will be to the tune of several hundred milliseconds. This is significant enough for sound to be heard as echo. Though it is possible to minimize this latency so that the returned sound is not heard as echo (but is instead heard as side-tone, which is acceptable and in some scenarios even desired), such an involved solution will definitely be a custom one. Custom solutions will put serious restrictions on scalability of the product in terms of feature and performance.

A simple way to minimize Acoustic Echo is via the inclusion of an Acoustic Echo Cancellation (AEC) module in the audio path. Though software implementation of AEC is available under ARM, it comes at a price – it demands significant CPU clocks. Further tuning of such an AEC module for specific deployment condition will further require expert-level-customization in the firmware. This approach therefore is not attractive while developing a scalable, feature-rich product.

A solution that does not restrict the scalability of the product and can be easily configured for the various deployment scenarios is the hardware based AEC. IP Video Door System incorporates such a hardware based AEC module in the path of audio circuitry within the Entry Unit. This eliminates the echo and thus supports all the following hands free conversations scenarios including Double Talk:

- Single cast – Alternatively known as Single Talk - where one person talks and the person at the other end listens
- Half-duplex – Also known as Alternate Talk – where both people talk, but alternately, one by one
- Full-Duplex – Known as Double Talk – where both people talk simultaneously

This AEC chipset with powerful AEC capability (65 dB) including non-linear echo cancellation has excellent inbuilt noise suppression and side-tone attenuation (25 to 35dB).

Design Approach ▶▶▶

The design of the IP Video Door phone system keeps in mind the need to include and exclude features for various deployments scenarios. The modular design of the system makes use of software components that are configurable to ensure ease of maintenance, configuration and upgradability.

Wherever feasible, executable binaries support multiple deployment scenarios through run time configuration. The source code of the software modules implements all the required features and supports run time configuration by means of input options.

System Environment

Hardware Environment:

Both Entry Unit and Apartment Units are powered by ARM based SoC from TI – DM365 running at 270/300MHz. They use Ethernet for inter-connectivity and are powered by PoE switches. They can be customized to be powered by an external power adapter.

Software Environment:

Both Entry Unit and Apartment Units run on MontaVista Linux Pro 5.0.0 ported for TI DaVinci DM365 platform.

The Linux kernel version is 2.6.18.

The GUI subsystem is based on QT 4.6 which is an open-source distribution under support by Nokia.

About Calsoft Labs

Calsoft Labs provides specialized concept to market Product Engineering services to product and technology companies in select market segments. Our target markets include Automotive, Consumer Electronics, Media, Networking, Storage and Independent Software Vendors (ISVs). Calsoft Labs delivers unmatched business value to its customers through a combination of process excellence, reusable frameworks and technology innovation.

Calsoft Labs is a wholly owned subsidiary of ALTEN. Set up in 1988, ALTEN is a European leader in Engineering and Technology Consulting (ETC) with 15,000 employees in over 14 countries worldwide.

USA
2903, Bunker Hill Lane, Suite 107, Santa Clara, CA 95054
Phone : +1 925 249 3000 - Fax : +1 925 249 3031

INDIA
196, Bannerghatta Road, Arekere Circle Bangalore - 560076
Phone : +91 80 4034 3000 - Fax : +91 80 4034 3111

FRANCE
40 avenue André Morizet, 92514 Boulogne-Billancourt, France
Phone : +33 (0)1 46 08 70 00 - Fax : +33 (0)1 46 08 70 10