Future Mobile Phones - Complex Design Challenges from an Embedded Systems Perspective

Fredrik Dahlgren

Ericsson Mobile Communications AB
SE-221 83 Lund, Sweden
fredrik.dahlgren@ecs.ericsson.se

1 INTRODUCTION

In the coming few years, three trends for mobile phones are seen. First, new communication standards provide packet-switched high-bandwidth wireless communication globally. Second, user applications are going from simple phone-administration and text-messaging to advanced multimedia and entertainment applications. Finally, because of Internet access and security support, the phones have the potential to move into a financial instrument for economical transactions. This talk will cover mainly the two former trends, and describe key complexity issues and design challenges.

This extended abstract is organized as follows. In Sections 2 and 3, the concept of Future Mobile Phones will be described; Section 2 covers communication capabilities while Section 3 points out some emerging user application capabilities. Section 4 describes and discusses complexity issues and some design challenges, and Section 5 provides a summary. Overall, this is intended as an introduction to the keynote presentation and an area overview, and will not cover the issues in great detail or propose technical solutions. The focus is from the digital system architecture's perspective, and does not cover challenges in the areas of radio, antenna, or analog/digital integration.

2 FUTURE COMMUNICATION CAPABILITIES

Today's cellular communication standards are digital and include GSM, TDMA, and cdmaOne, and are referred to as the 2nd generation (the first generation being the various analog standards). Their evolution into 3G (3rd generation) is shown in Figure 1. The presentation will mainly cover the european and asian systems, including the GSM evolution.

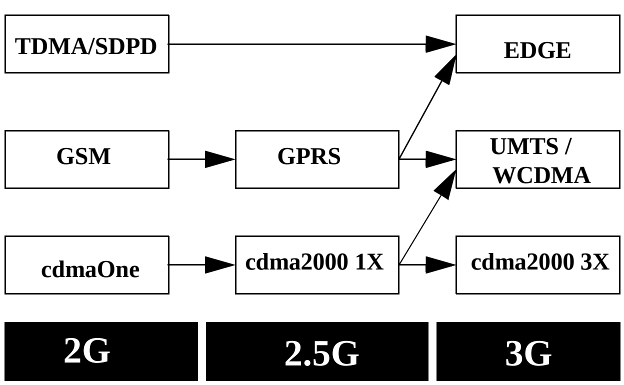

Figure 1: The evolution to 3rd generation (3G) cellular standards.The first step into 3G is often referred to as 2.5G.

GSM provides a circuit switched connection at 9.6 kbps. Being a multislot version of GSM, the HSCSD (High-Speed Circuit Switched Data) is a high-speed version and provides up to 57.6 kbps, circuit switched, when 6 slots are supported. GPRS (General Packet Radio Service) is a packet-switched extension to GSM, and provides up to 115 kbps. GPRS has recently been introduced in several european countries.

EDGE is an evolution of GSM that uses a new modulation technique, and brings the peak data rate to 384 kbps (packet switched). The same technique is used in the evolution of TDMA (primarily used in North America).

UMTS (Universal Mobile Telecommunication System), currently under development, is a standard that will be used world-wide[1]. It is based on WCDMA (Wideband CDMA), and will support up to 2 Mbps transfer rates (packet switched), and simultaneous circuit-switched and packet-switched services.

For close-range wireless connections, Bluetooth [1] has recently been launched. Bluetooth uses the globally available and license-free spectrum at 2.4 - 2.5 GHz. The data rate supported is either up to 433 kbps up- and downlink (symmetric), or 723 kbps in one direction and 57.6 kbps in the other. Range is either 10 or 100 meters (two different versions).

Besides communication with multiple cellular standards (for global coverage) and Bluetooth, the phones are likely to support IrDA, USB, and potentially other wired techniques and standards.

Looking ahead, there are efforts studying new modulation techniques for UMTS for even higher data rates. It is, however, unclear whether these will be targeting small phones and PDAs (Personal Digital Assistants) short term. In addition, there are also efforts going on studying and defining new versions of Bluetooth for higher data rates (some 10x improvements). Finally, even though laptops are likely to have in-built support for W-LAN (wireless LAN), a similar business case for small, form-factor sensitive phones remains to be seen.

3 FUTURE APPLICATION CAPABILITIES

Today's phones have rather poor user application capabilities, and besides speech and communication the functionalities are simple. This is soon to be changed. Although the form-factor limits the phones to a substantially smaller display, there are many applications, or types of applications, that will co-exist for both desktops and PDAs. This, however, does not imply that the phones will support only a subset of the desktop's functionalities. While a desktop has a substantially larger display and higher (or at least cheaper) communication bandwidth, the phone is always with you and will provide a plethora of applications related to positioning, E-commerce (electronic commerce), and lets the user be continuously in touch with people and information. Therefore, some applications will be more useful on a phone than on a desktop.

[1] A version of UMTS / WCDMA is being launched in Japan during 2001 which differs somewhat from the international 3GPP standard.

Proceedings of the Seventh International Conference on Engineering of Complex Computer Systems (ICECCS'01)
1050-4729/01 $10.00

Recently, a number of user applications have been launched for PDAs and phones, including video, still image support, high-quality audio, speech recognition, and games. In the future, more advanced versions of these will be available, but their capabilities will vary depending on price and form-factor of the products.

For video, examples of services include decode of video clip that has earlier been stored, or streaming video decode of a news or sports broadcast. Videoconferencing or videophone is a service where two parties support both video encode and decode. Video encode can also be used for capturing simple clips or as a replacement of a video camera. Common standards for video include MPEG-4 [2][3] and ITU's H.263 [4][5], but also other formats that are common on Internet will be supported.

Still images can be encoded and decoded to / from a number of different formats, such as JPEG and the emerging JPEG2000 standard (ITU-T Rec. T.800, ISO/IEC 15444-1) [6], the latter being based on wavelet transforms. High-quality audio typically refers to formats for compressed digital music in stereo, such as MP3. There exist several different compression formats, and one important aspect is currently support for protecting the contents so that only authorized people can decompress and listen to the music, such as Watermarking techniques [7].

Several phone manufacturers provide support for name dialing with Automatic Speech Recognition (ASR). On the other hand, there exist ASR software for advanced dictation on desktop computers. In this context, it is important to make a distinction between noise-robust ASRs [8], suitable for phones and PDAs that are often used on streets, in shopping malls, or in cars, and large-vocabulary ASRs [9] developed for silent office environments. However, looking ahead, phones that lack suitable keyboard capabilities are likely to be a more natural target for very advanced ASR support also including dictation.

Games are often pointed out as one of the most important applications for future phones. Depending on form-factor of the phone (in this case: display size and resolution), advanced graphics will be one key characteristic of the platform. It is likely that we will see three dimensional graphics support in games rather soon. Such advanced graphics will also be used at wireless shopping, to provide a virtual hands-on of various products.

Besides above applications, there are several other functionalities of the phones that will be critical:

- Support for cryptography for e-commerce and integrity
- Java, and other platform-independent execution environments
- Support for downloading of applications over the net, or to program the phone with some own applications or functions
- Upgrading existing functionalities over the net
- Browsing the normal Internet as well as support for WAP (Wireless Application Protocol) [10]
- Reading and editing documents, slide-presentations, etc.
- Positioning (e.g. GPS) and related services and applications
- Security control, such as biometric fingerprint recognition

Applications such as Speech (new codecs with improved audio quality are emerging), Email, calendar, phone-book, multimedia messaging services (MMS - the evolution of short message service, SMS, extended with multimedia contents) are introduced already now, and will be taken for granted by the user.

A large number of other applications can be described and discussed, but it is out of the scoop of this paper provide an extensive list.

4 Complex Design Challenges

4.1 Complexity Issues and Design Trade-offs

Looking briefly at the system, some general complexity issues can be identified. There are applications and functions with hard real-time requirements (e.g. the physical layer of the wireless communication), as well as with soft real-time, quality-of-service, requirements (e.g. higher communication layers, most multimedia-related user applications). The functional complexity of the emerging communication standards (e.g. UMTS) is very high, partly since the development takes place in parallel with the standardization process, and partly because many functionalities that are optional from the system network side are mandatory for the terminals (the terminals shall operate globally, with a large number of different operators and system infrastructure providers). While parts of the system must be secure and reliable, others have demands on openness and relatively high execution performance. With the adaptation to the Internet and emerging services, the phones will need to support new decoders for multimedia contents, although at the same time be small, cheap, and consume a minimum of energy.

From a hardware perspective, there are some clear design trade-offs.

- A high processing performance demands fast and advanced CPUs and DSPs, much on-chip memory, and enough and fast off-chip memory.
- The form-factor limits the size of chips and total number of components.
- Cost-effectiveness demands small chips (e.g. little on-chip memory, not as aggressive processor designs), little off-chip memory (or cheaper / slower memory types).
- A small energy consumption demands ample on-chip memory (fewer off-chip accesses), energy-effective off-chip memory, and hardware support for things that are expensive to do in software. Dynamic as well as static power dissipation should be kept on a minimum. CPUs and DSPs must be energy efficient.
- In addition, the wide software base and strict requirements on time-to-market call for flexible solutions.

Overall, there are several conflicting requirements, with several important design trade-offs. In the presentation, these are described and discussed in more detail.

From a software perspective, a phone will evolve from a completely embedded system to a system where parts are embedded and other parts are open. While being compliant to security requirements and operator-driven requirements, the platform should also support most upcoming trends initiated by users during the lifetime of the product. From a software architecture perspective, there are several trade-offs to consider, also baring in mind that the amount of memory is a limited resource and one of the more costly resources of the phone. Given the lack of a few dominant software suppliers and one dominating operating system, as opposed to desktop systems, this will certainly be an exciting area over the upcoming decade.

4.2 Design Challenges

One true design challenge is to provide ample processing performance that meets the requirements, at a minimum average power dissipation, with few and small (cheap) chips. Below, some challenges are pointed out and motivated. The talk will cover this topic in more detail.

For the communication part, which is subject to type approval (certified to be used in a cellular system), a correct functionality is extremely important. Risks must be minimized. However, for high-speed communication (e.g. UMTS), certain parts of the baseband signal processing and control are both complex and demanding from a performance / power perspective, and an efficient design is still important. This is clearly a hardware / software partitioning trade-off, and real-time requirements must be secured by design. Since the communication functionalities is directed by standards, there are certain major design steps (from 2G to 2.5G to 3G), while the evolution is clearly incremental inbetween. While extremely challenging, the design evolution follows relatively long-term roadmaps, although in practice the standardization work is often delayed and from a market perspective the products should come close after the standards have been set.

For new standards, the hardware and software are co-developed, often in parallel with the standardization work, and with strict real-time requirements on parts of the system. To keep the design cycle short, there is a need for suitable tools and methodologies supporting efficient hardware / software co-design and verification - not only on the silicon-near level but also for software on a higher abstraction level (where millions of instructions need being executed).

Overall, for the communication parts, one of many design challenges is to define an architecture that is relatively predictable while still being small, power efficient, and compatible with state-of-art compilers, tools and software.

Requirements on user applications change more rapidly, and are subject to market trends. Although some long-term roadmaps are set up, there is a great need for fast adaptability and flexibility.

From a hardware perspective, the design challenge is to find processor- and memory system solutions that provide enough performance for the worst case scenario (that shall be supported), be small, average a minimum of energy consumption, and be compliant with software development and the existing software base (internal and external to the phone manufacturer). In addition, the hardware architecture should be scalable, to enable a maximum of hardware and software re-use across a product range.

In terms of processing capabilities, there is a plethora of possibilities with various CPUs, DSPs, and hardware accelerators. For the most demanding applications, hardware / software partitioning is an important design challenge, also taking into account the flexibility needed by the user to enable new applications.

For the memory system, there exist many different technologies possible for both on-chip and off-chip memories. For on-chip memories, there are fast but large SRAM (single- or dual-ported, cells optimized for speed or for area, etc.), embedded high-density DRAM, embedded high-density non-volatile memory (e.g. Flash), and masked ROM. Off-chip memory technologies include SRAM, DRAM, and non-volatile memories (e.g. Flash). There are various technologies for non-volatile memory that differ in cost, speed, power, and density, ranging from burst-mode fast Flash memories with a wide interface, to Flash cards using a serial interface. While on-chip memories can be configured and optimized according to needs, off-chip memories are dictated by global demand/supply rules for any type of system. Even though an on-chip memory is faster, consumes less power, and allows for a potentially lower pin-count and fewer components, the cost per byte of on-chip memory of often larger, and embedded DRAM and non-volatile memories leads to more mask layers which further increases the cost of the chip.

The characteristics of the application program execution directs which processor capabilities and which memory system that are best suited. These should be dimensioned after the most demanding applications and user scenarios, although scalability issues from entry-level up to high-end products must be considered. However, the width of the software base, the operating systems to be supported, and the planned sources of the software, must direct the base of the system (CPUs and DSPs, etc.).

5 Summary

Mobile phones are becoming extremely complex systems despite a small physical size. Functionalities include multi-standard high-speed wireless communication capabilities, positioning, financial-level security support, and aggressive multimedia and business applications. The phones are therefore becoming an instrument for business, pleasure, and commerce. The small physical size increases the design challenge.

The software system is becoming increasingly complex, and includes hard as well as soft real-time requirements, and secure parts as well as openness. The hardware system is challenging with conflicting demands on performance, cost, silicon area, and size as well as type of on- and off-chip memories.

For certain functionalities of the system, hardware and software are co-designed, and the requirements include keeping the risks low while still targeting an efficient design. A critical area is therefore tools and methodologies supporting hardware / software co-design and verification. For certain other features, the architecture (hardware *and* software) of the system must be both flexible and scalable.

6 References

[1] K. Diefendorff, "Bluetooth Creates Personal Wireless Network", Microprocessor Report, June 1, 1998.

[2] ISO/IEC 14496, "Information Technology - Coding of Audio-visual Objects", 1999.

[3] L. Chiariglione, "MPEG and Multimedia Communication", *IEEE Transactions on Circuits and Systems for Video Technology*, 7(1):5-18, February 1997, pp.58-64.

[4] ITU-T Recommendation H.263 "Video Coding for Low Bit-rate Communication", 1996.

[5] C.-W. Ku et al., "Investigation of a Visual Telephone Prototyping on Personal Computers", *IEEE Transactions on Consumer Electronics*, 42(3):750-758, August 1996.

[6] C. Christopoulos, A. Skodras, and T. Ebrahimi, "The JPEG2000 Still Image Coding System", *IEEE Transactions on Consumer Electronics*, 46(4):1103-1127, Nov. 2000.

[7] F.A.P. Peticolas, "Watermarking Schemes Evaluation", *IEEE Signal Processing Magazine*, September 2000.

[8] R. Mammone, X. Zhang, and R. Ramachandran, "Robust Speaker Recognition - A Feature-Based Approach", *IEEE Signal Processing Magazine*, September 1996, pp.58-71.

[9] S. Young, "A Review of Large-Vocabulary Continuous-Speech Recognition", *IEEE Signal Processing Magazine*, September 1996, pp.45-57.

[10] WAP Forum, http://www.wapforum.org/