Evolving Cellular Handset Architectures but a Continuing, Insatiable Desire for DSP MIPS

Michael L. McMahan *Texas Instruments Incorporated*

ABSTRACT

The year is 2005. Wireless technology has become so pervasive that kindergarten children and family pets are tracked with GPS receivers; businessmen work in virtual offices with high-speed wireless connections; shopping, stock trading, and ticket purchases are all done routinely using wireless terminals that are worn—not carried. Utopia—or your worst nightmare—whatever your perspective, this is a future which is just over the horizon. It follows down a path which began with Hertz and Marconi experimenting with radio transmission through the "ether" in the late 19th century and continues today with the explosion of wireless products and services.

Contents

List of Figures

List of Tables

This document was originally published in the Texas Instruments Technical Journal, January-March 2000, Volume 17, Number 1.

1 Introduction

Twenty years ago, few would have predicted the revolution which resulted from the development of the personal computer in the late 1970s. It is easy to forget that cellular telephony is an even more recent phenomenon. The first commercial cellular system in the United States was turned on in Chicago in 1983. As recently as 1988, George Calhoun, in his book, *Digital Cellular Radio*, talked about the reasons for the failure of cellular technology in the marketplace. Today, a decade later, cellular telephony is one of the largest, fastest growing markets in the high-tech marketplace. In the last two years, the worldwide consumption of cellular handsets has tripled from 48M units in 1996 to 153M units in 1998. Cellular penetration has experienced a compounded growth rate of approximately 60%–70%. This paper gives a short tutorial background on cellular standards, describing the current evolution of second to third generation systems, and discusses how handset architectures, in general, and digital signal processor (DSP) technology, in particular, must evolve to meet the needs of these future systems.

2 Cellular Standards

Figure 1 describes the evolution of cellular standards from the first generation (1G) of many incompatible analog standards to a third generation (3G) that promises to be all digital, but almost as confusing.

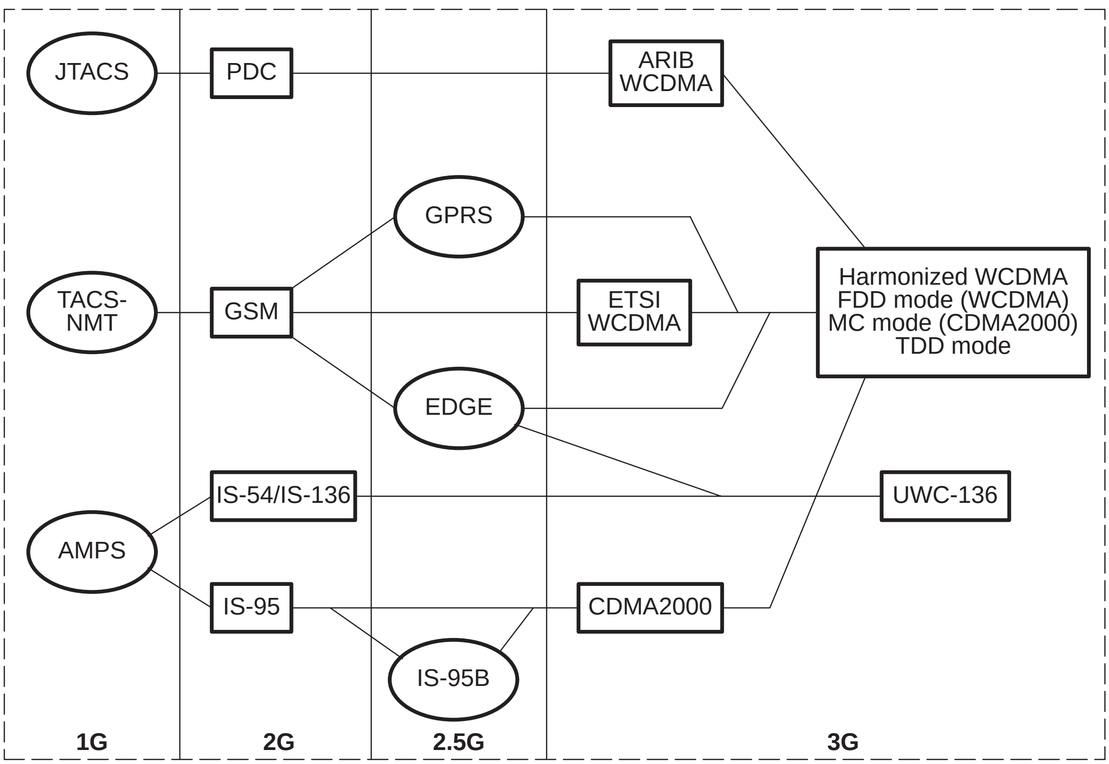

Figure 1. Evolution of Cellular Standards

Analog cellular systems use a frequency division, multiple access scheme to support multiple users within a finite geographical area (see Figure 2a). The AMPS standard divides the allocated spectrum into 30-kHz channels—one user per channel. TACS, NMT, and JTACS used a 25-kHz spacing. Within a single channel, all of these standards use frequency modulation (FM) to represent the speech waveform. AMPS, for example, modulates the carrier up to ±12 kHz to represent the amplitude of the waveform.

The 1990s saw the emergence of the second generation (2G) of cellular systems based on digital technology. Digital technology was enabled by the emergence of cost-effective, high-performance DSPs—particularly the Texas Instruments (TI) TMS320C54x. Initially, the transition was driven by cost. Because digital systems were able to multiplex more users in a given block of frequencies, it was cheaper to support a digital user than an analog one. Moreover, digital technology could support a wide range of services beyond voice telephony. After all, bits are bits. It does not matter whether those bits represent voices or stock quotes. This fact gave the service provider the ability to begin to differentiate his service from his competitors. From the user perspective, the key advantage of digital standards is increased battery life. Analog phones have to constantly monitor the channel to check for incoming phone calls. This burns precious power. Digital handsets on the other hand, support the ability to enter into a sleep mode, which consumes very little power, and then wake up periodically to check the channel for incoming calls. The standby specification for analog phones is measured in hours. It is not uncommon for modern digital cell phones to exhibit standby times of days or even weeks. The second advantage of digital technology for the consumer is security. Analog cellular telephony is notoriously easy to monitor or spoof. Digital technology is amenable to incorporating the latest encryption technology and is consequently much more secure.

Second-generation cellular telephones are classified based on the technique which is used to multiplex multiple users onto a single carrier frequency (see Figure 2b, Figure 2c, and Table 1).

GSM and IS-136 are both examples of time division multiple access (TDMA) systems. This technology allows “n” users to share a single carrier frequency by giving them unrestricted access during a fraction (n/T) of each frame (see Figure 2b). IS-136 uses a channel width of 30 kHz (compatible with AMPS), divides time into 40-ms frames, and shares that time among 3 users. GSM uses a 200-kHz channel spacing, divides time into 4.62-ms frames, and divides each frame into 8 different user slots.

IS-95 is based on a spread spectrum technology call Code Division Multiple Access (CDMA). This standard divides the spectrum into 1.25-MHz channels. The standard supports up to 64 users (though practically only about 20 can operate simultaneously) in each channel. Separation among these users is accomplished by spreading the data signal of each user with a unique orthogonal digital code (see Figure 2c). The receiver then correlates the received bit stream with that digital code to extract a specific user’s data. Since the spreading codes are orthogonal, users whose data is spread using a different code appear as white noise. Consequently, as additional users are added, the signal-to-noise ratio (SNR) experienced by any individual receiver will rise and the quality of a call (as measured by bit error rate (BER)) will gradually degrade. This is different from TDMA systems, which have a hard limit on the number of users who can simultaneously use a specific cell. When that number is exceeded, new calls are blocked.

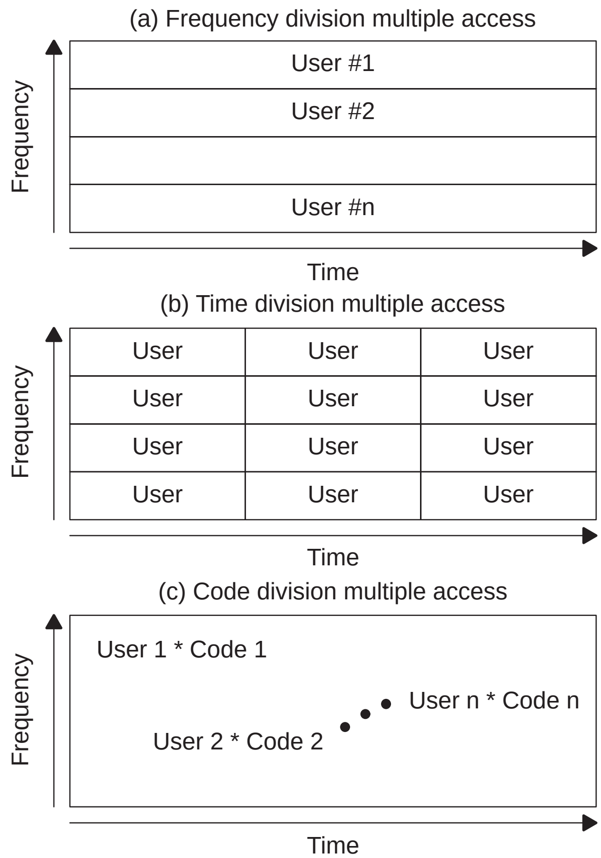

Figure 2. Multiaccess Technologies

Table 1. 2G Cellular Standards

	GSM	IS-136	IS-95	PDC
Frequency (MHz)	GSM900	Cellular	Cellular	
Forward	935-960	869-894	869-894	940-956
Reverse	890-915	824-849	824-849	810-826
Forward	GSM1800			1429-1453
Reverse	1805-1880			1477-1501
	1710-1785			
Forward	GSM1900	PCS	PCS	
Reverse	1930-1990	1930-1990	1930-1990	
	1850-1910	1850-1910	1850-1910	
Carrier Spacing (kHz)	200	30	1250	25
Access Method	FDMA/TDMA	FDMA/TDMA	FDMA/CDMA	FDMA/TDMA
Users/Carrier	8	3	~20	3
Duplexing	FDD	FDD	FDD	FDD
Modulation	GMSK	π/4 DQPSK	BPSK/ODQPSK	π/4 DQPSK
Speech Data Rate	13 Kbps	7.95 Kbps	13 Kbps (variable)	6.7 Kbps

Today, at the end of the 20th century, the transition from analog to digital is largely complete. The biggest trend today is from voice-centric to data-centric systems. Remember that one of the key advantages of digital technology is its ability to support a wide variety of services. However, 1G and even 2G telephones are relatively poor data devices because of their limited bandwidth. Most digital standards only support about 9.6–14.4 Kbps. While this may be sufficient for services like GSM's Short Message Service (SMS) or very basic data services like stock quotes, sophisticated services involving internet access need more. That is what is driving the next transition—to 2.5G and even 3G standards.

Each of the 2G standards has some kind of plan to support higher-bandwidth services. Most of these beong to a class of systems which has been categorized as 2.5G. Most of these also include a transition from circuit-switched to packet-based communications. The value of packet-switching is that it gives the consumer an "always-on" service, which eliminates the long call setup time that is built into any circuit-switched connection. Generalized Packet Radio Services (GPRS) is probably the best and most advanced example. GPRS combines GSM slots to provide as much as 115 Kbps in packet-switched, always-on bandwidth. Another 2.5G standard is called Enhanced Data rates for GSM Evolution (EDGE). EDGE will use higher-order modulation (either GMSK or 8 PSK) to support data rates of up to 384 Kbps—basically the same rates as advertised for 3G pedestrian-class applications. Significant GPRS rollout should begin in 2000. Deployment of EDGE is much less certain, but is probably still several years away.

Third-generation standards were developed specifically to support high-bandwidth data services. The dominant 3G standard, Wideband CDMA (WCDMA), is based on the same class of technology as IS-95. At a high level, the major difference is that IS-95 spreads the signal to 1.25 MHz with a chip rate of 1.2288 Mcps, whereas WCDMA spreads to 5 MHz with a chip rate of 3.84 Mcps. WCDMA will support data rates of 144 Kbps for high mobility applications, 384 Kbps for more pedestrian-class mobility, and 2Mbps for fixed, office environments. The whole thrust for both 2.5G and 3G cellular system is pointed at data applications. That fact is important to keep in mind as we look at the implementation architectures that are used to deliver today's 2G handset and which will be required to support 2.5G and 3G in the future. 3G deployment will begin in Japan in late 2001 with a significant ramp in 2002.

3 Handset Architectures Today

Figure 3 is a functional block diagram of a generic 2G voice-centric cellular phone. The architecture which implements this system is based on two processors (generally integrated onto a single integrated circuit).

The heart of all digital cellular phones is a DSP. The dominant DSP in this market is the Texas Instruments TMS320C54x, which can be found in about 65% of modern cell phones. This processor is responsible for modulating and demodulating the data stream, coding and decoding to maintain the robustness of the transmission in the face of transmission bit errors, encrypting and decrypting for security, and compressing and decompressing the speech signal. In early 2G TDMA phones these functions could be accomplished with 30–50 DSP MIPS. As vocoders have become more complex and as data rates have risen in 2.5G phones, the total DSP load has risen past 100 MIPS. The CDMA standard requires a somewhat different functional partitioning because of the data rates generated by spreading. While the DSP can still be used to process at the basic data rate (functions like forward error correction, encryption, or voice compression), ASIC hardware operating under the control of the DSP must be used to process and modulate/demodulate the spread signal.

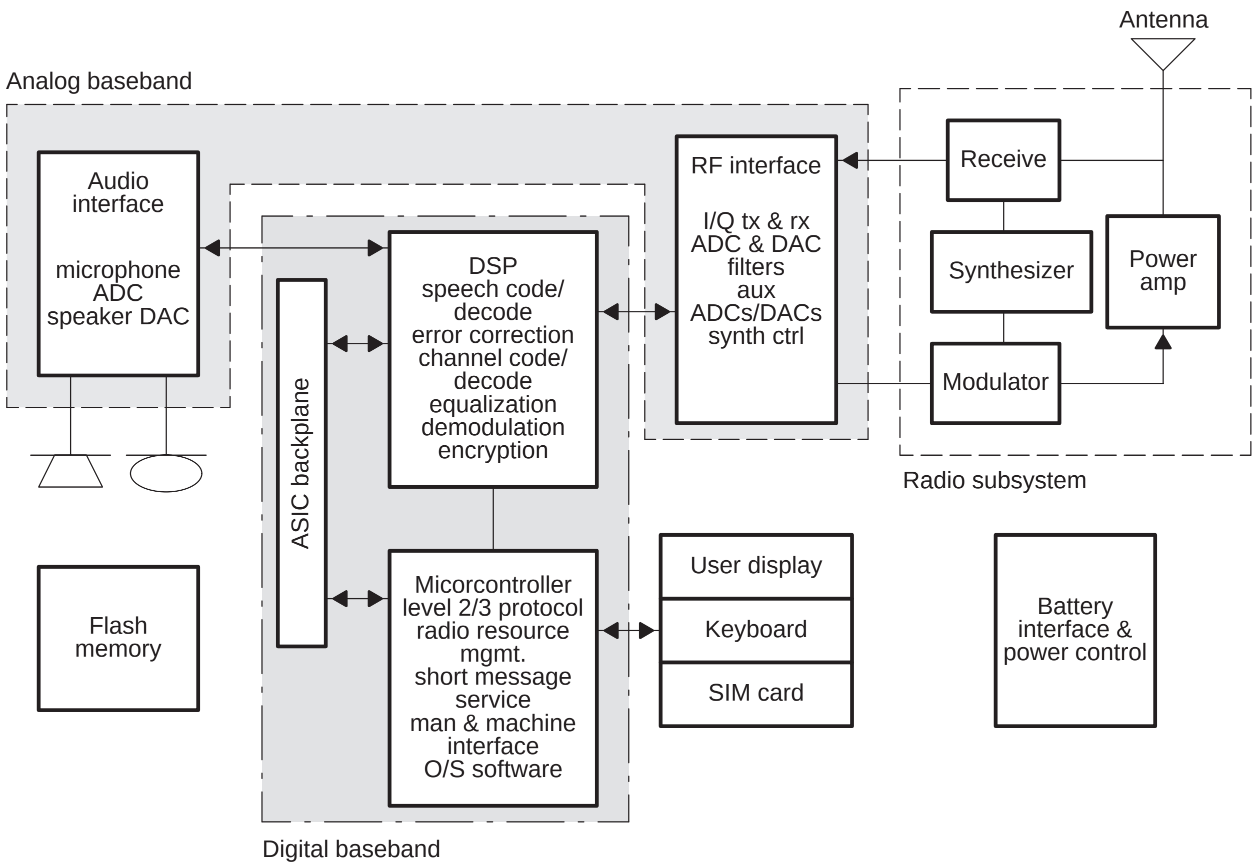

Figure 3. Generic 2G TDMA Baseband Architecture

The second processor that is found in a 2G phone is a general-purpose microprocessor which supports the user interface of the handset and which handles the upper layers of the communications protocol stack. In most modern 2G phones, this processor is implemented with a 32-bit RISC engine (generally operating in a 16-bit mode to conserve memory). For a voice-centric handset, the functions implemented by this RISC processor require less than 10 MIPS. Even for many data-intensive 2.5G phones, the MIPS requirement of this processor will be minimal—probably less than 40 MIPS.

4 And in the Future

Future handsets will change along multiple dimensions. First, radio architectures are evolving in directions that could dramatically change the partitioning illustrated in Figure 3. The goal is to support more flexibility in the radio frontend and to reduce the number of components required. In many cases, this means that some of the functionality that is implemented with analog components today will shift to the digital signal processing domain. For example, current narrowband analog receivers could be replaced with wideband analog frontends that use programmable digital techniques to filter and select channels. Other examples include using adaptive digital techniques to linearize transmit power amplifiers or to eliminate the DC offset in direct conversion receivers.

On a different dimension, as data applications become more important, the nature of a mobile handset will change. Today, a handset is the perfect example of a closed, static, embedded system. Its function is defined when the phone is manufactured. Although it may support a few simple applications beyond voice communications (like SMS or even limited information services), the system exists to implement a voice communications channel—a fixed, real-time DSP-intensive task. In the future, it seems likely that the market for mobile handsets will fragment. We will probably always have a segment of the market for the traditional, voice-centric mobile phone. However, we will also have hybrid devices where the modem communications function is a necessary component, but one or more data applications may well be the reason the device is purchased. In some cases, these applications may be fixed, but given the advent of the Internet, it seems more likely that applications will be downloadable and will change at the whim of the consumer. Today, I might choose to download an MP3 player. Tomorrow, I might want my phone to support a video conference or act as a GPS navigator. In a sense, the handset has taken on a very PC-like character. It has become an applications platform (see Figure 4). What are the implications for handset architectures?

Figure 4. Mobile Data Applications

As long as the processing load of a cellular handset was fixed, it was relatively easy to size and configure the processing resources required to implement the communications modem. In this new environment, the requirement for a cellular modem is unchanged (though the data rates may be substantially higher). However, the applications requirements are potentially unbounded. The question is how to best add processing resources to handle these additional tasks. The answer is complicated by the fact that applications—like the modem—will have both control and DSP requirements. As in most complicated engineering problems, there is no single best answer, but the major issues are clear:

1. How do you supply the processing resources necessary to handle an open-ended, unspecified, and potentially unbounded load ?

   Because the platform will need to support application flexibility, its basic architecture will be based on one or more programmable processors. This was true in 2G phones because of ever-changing communications standards and time-to-market issues. This will continue to be true in the future, but the nature of these processing elements and the software they execute will change.

Second-generation handsets were very DSP-centric. That fact will remain true in the future. However, because of the need to support sophisticated applications, 2.5G and 3G handsets will require a more balanced set of processing resources. Either the current RISC engine must be enhanced or a separate RISC platform must be added to the architecture. This processor will need to operate at >100 MIPS and will need to incorporate features to allow it to support much more significant operating systems (OSs) than we have seen in traditional cell phones. Applications-centric OSs such as EPOC from Symbian, Windows™ CE from Microsoft™, or the PalmOS™ from 3COM™ will become the order of the day. The software environment of the DSP will also have to change in order to support downloadable applications (which is likely considering the spectrum of applications shown in Figure 4). Since the DSP resource will have to support the dramatically increased communications load imposed by high-bandwidth 3G standards (>200 MIPS) and also satisfy the demands of a range of applications-specific DSP functions (speech recognition, image/video coding and decoding, etc.) which themselves will drive a significant MIPS load, it is possible that these future handsets will require more than one DSP.

2. How do you provide substantially increased processing resources without decreasing battery life?

   The consumer is quite demanding. Now that he has experienced operating battery life in excess of 4 hours and standby times measured in days or weeks, he will never be willing to accept anything less. The challenge will be to supply an order-of-magnitude increase in processing power without no decrease in battery life. Part of this challenge will be met with advances in silicon technology, but part will have to be achieved with architectural innovation. Examples could include:

   a. More care in matching processing resources to the type of algorithm required. This implies a mix of general-purpose RISC and DSP resources. Although many simple DSP tasks can be accomplished using a RISC instruction set, it is often more power efficient to use a DSP.
   b. Innovative use of attached, programmable coprocessors. These can often supply more power-efficient execution than a fully programmable DSP, yet retain much of the flexibility.
   c. More extensive use of multiprocessor architectures to minimize clock rates required to accomplish task. This would have the further benefit of allowing unused functions to be turned off to minimize standby power consumption. Such architectures will also drive a need for new operating system and software technology.

3. How do you integrate the requirements of hard real-time communications tasks with the requirements of non- or soft-real-time applications ?

   The answer may well be that you do not. In order to maximize system reliability and minimize the software development complexity of any platform, it may be necessary to segregate real-time and non-real-time tasks. None of the OS environments mentioned above are particularly well-suited for supporting the predictable, deterministic requirements of high-speed wireless modems. Although all of these software platforms are moving to incorporate real-time features, it will take time for designers to become comfortable with their ability to achieve the reliability and performance which these systems require.

5 Conclusions

Future data-centric, 2.5G and 3G mobile handsets will require a new approach to the architectures which implement those devices. Care must be taken to consider application and user interface performance, system-level power consumption, hardware costs, software complexity, and time to market. The winning system will likely incorporate multiple processors tuned to fit the tasks for which they were designed. You will see high-end RISCs, which implement protocol stacks, user interfaces, high-end OSs, and downloadable applications. You will also see programmable DSPs and coprocessors that can provide power-efficient media processing and support needed application flexibility and upgradeability. The biggest open question is the extent to which the applications and communications functions of these devices can be effectively combined into one programming environment. In the near future, it is likely that they will be separate.

6 References

1. U. Amin and S. Chennakeshu, “The Evolution of TDMA to 3G,” guest editorial, *IEEE Personal Communications*, Vol. 6, No. 3, Jun. 1999, pp. 6–7.
2. P. Bahl, “Supporting Digital Video in a Managed Wireless Network,” *IEEE Communications*, Vol. 36, No. 6, Jun 1998, pp. 94–102.
3. G. Calhoun, *Digital Cellular Radio*, Artech House, 1988, pp. 1–54.
4. A. Gatherer, T. Stetzler, M. McMahan, and E. Auslander, “DSP-Based Architectures for Mobile Communications: Past, Present, and Future,” to be published in *IEEE Communications*, Jan. 2000.
5. U. Ko, M. McMahan, and E. Auslander, “DSP for the Third Generation Wireless Communications,” to be published in *ICCD ’99*.
6. Z. Kostic and S. Seetharaman, “Digital Signal Processors in Cellular Radio Communications,” *IEEE Communications*, Vol. 35, No. 12, Dec. 1997, pp. 22–43.
7. M. McMahan, A. Khatzibadeh, and Pradeep Shah, *Wireless Systems and Technology Overview*, internal Texas Instruments white paper, 8/95.
8. M. Oliphant, “The Mobile Phone Meets the Internet,” *IEEE Spectrum*, Aug 1999, p. 20-28.
9. N. Sollenberger, N. Seshadri, and R. Cox, “The Evolution of IS-136 TDMA for Third-Generation Wireless Services,” *IEEE Personal Communications*, Vol. 6, No. 3, Jun. 1999, pp. 8–18.