

# DSP-Based Architectures for Mobile Communications: Past, Present and Future

*Alan Gatherer, Trudy Stetzler, Mike McMahan, and Edgar Auslander, Texas Instruments*

## ABSTRACT

Programmable DSPs are pervasive in the wireless handset market for digital cellular telephony. We present the argument that DSPs will continue to play a dominant, and in fact increasing, role in wireless communications devices by looking at the history of DSP use in digital telephony, examining the DSP-based solution options for today's standards, and looking at future trends in low-power DSPs.

## INTRODUCTION

Programmable digital signal processors (DSPs) are pervasive in the wireless handset market for digital cellular telephony. There have historically been two distinct approaches to the implementation of cellular handsets and infrastructure equipment. One approach emphasizes programmable DSPs for flexibility in the face of changing standards, while the other approach focuses on hard-wired application-specific integrated circuit (ASIC) techniques to improve implementation efficiency. The right answer has been and will likely continue to be some combination of both approaches.

In this article we present the argument from a pro-DSP perspective by looking at the history of DSP use in digital telephony, examining the DSP-based solution options for today's standards, and looking at future trends in low-power DSPs. We show that some very compelling arguments in favor of the unsuitability of DSPs for second-generation digital telephony turned out to be spectacularly wrong and that, if history is to teach us anything, it is that DSP use increases as a wireless communications standard matures. We summarize some of the up and coming applications for the new third-generation wireless personal assistants to show that, if anything, flexibility is becoming more of an issue, and therefore the programmability offered by DSPs is even more desirable. Since power is the greatest potential roadblock to increased DSP use, we summarize trends in power consumption and millions of instructions per second (MIPS), and we also describe the use of coprocessors as a complementary technology to today's DSP, allowing flexible solutions to third-generation wireless.

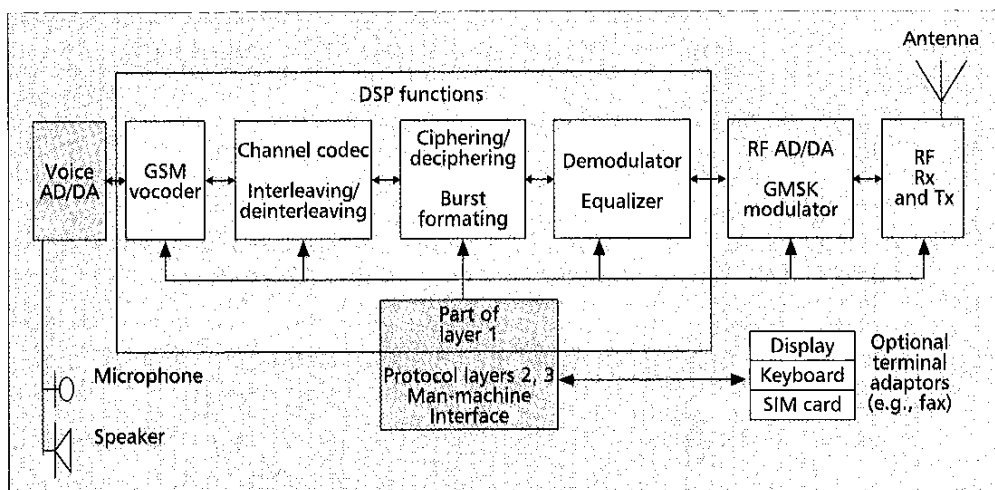
## A HISTORICAL PERSPECTIVE ON WIRELESS HANDSET ARCHITECTURES FOR GSM

As we mentioned in the introduction, there is a continuing debate over the role of DSPs in wireless communications. To provide a historical basis for our arguments, in this section we examine the case of Global System for Mobile Communications (GSM) evolution. The assumption is, of course, that third-generation cellular (3G) products will evolve in a similar manner to GSM, which is in itself debatable, but we believe that history does have some good points to make with respect to 3G.

A common functional block diagram of a GSM system is given in Fig. 1. We recognize a classical digital communication model with signal compression, error correction, encryption, modulation, and equalization. In the early days of GSM it was assumed that the low power requirement would mean that most of the phone would be implemented in ASIC. In what follows we show that the power difference between DSP and ASIC was not significant enough compared to other factors that drove GSM phone evolution.

The coder used in GSM phase 1 compressed the speech signal at 13 kb/s using the rectangular pulse excited linear predictive coding with long-term prediction (RPE-LTP) technique as per GSM 06-10 specification. The voice coder is the part of the architecture that most engineers agree should be done on a DSP, and in early designs the DSP was included mainly to do the vocoding. However, once the DSP was included a certain amount of "mission creep" started to occur. As DSPs became more powerful, they started to take on other physical layer 1 tasks until all the functions in the "DSP functions" box in Fig. 1 were included.

Flexibility was also important in the evolving standard. GSM phase 2 saw the introduction of half rate (HR) and enhanced full rate (EFR). HR was supposed to achieve further compression at a rate of 5.6 kb/s for the same subjective quality, but at the expense of an increased complexity, and EFR had to provide better audio qualities and better tandeming performance,



■ Figure 1. A functional block diagram of a GSM phone.

also at the expense of higher complexity, using an enhanced Vector-Sum Excited Linear Prediction (VSELP) algorithm. Along with these changes came changes in the implementation of the physical layer as better performance, cost, and power savings combinations were found. As a result, each generation of phone had a slightly different physical layer from the previous one, and upgrades to ASIC-based solutions became costly and difficult. Because DSPs were now being designed with low-power wireless applications in mind, the power savings to be had from ASIC implementation of DSP functions was not significant enough that system designers were willing to live with the lack of flexibility. After 1994, a single DSP was powerful enough to do all the DSP functions, making the argument for a DSP-only solution for the baseband even more compelling. To improve system power consumption and board space, several DSP such as the Motorola 56652 [1] and TI's Digital Baseband Platform [2] integrate a RISC microcontroller to handle the protocol and man-machine interface tasks to free the DSP for communication algorithm tasks. The presently most popular partitioning of GSM is as shown in Fig. 2.

It is also true that as GSM phones have evolved, they have gradually moved beyond the simple phone function, and this has led to an increase in the fraction of the DSP MIPS used by something other than physical layer 1. This evolution is shown in Fig. 3. With the advent of wireless data applications and the increased bandwidth of 3G, we expect this trend to accelerate. We discuss this more later.

To summarize, despite the initial belief that a full ASIC implementation would be the norm for GSM, this turned out to be far from the case for the following reasons:

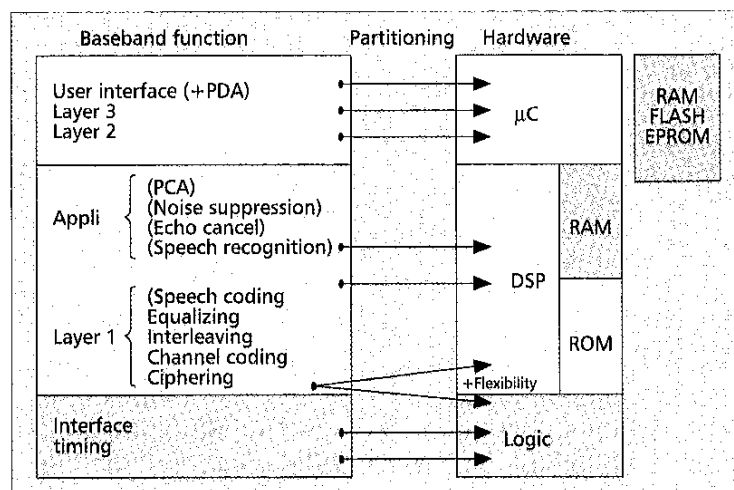
- Making an ASIC vocoder was like replicating available commercial DSP architectures. This provided the initial hook to get the DSP into the system.
- The product life cycle shortened from 2.5 years to 1 year thanks to technology and understanding, but also thanks to the phone becoming a personal fashion statement.

DSPs have a significant advantage over ASIC when the product life cycle is short.

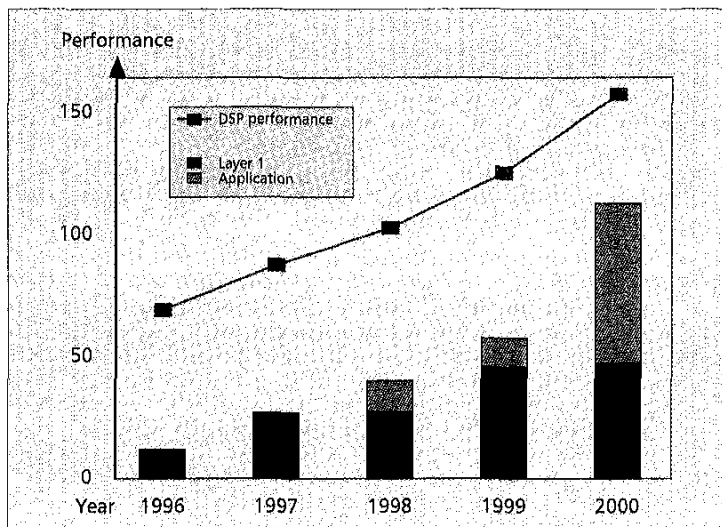
- Different worldwide standards related to GSM and the need for product families addressing different market segments called for a platform-based architecture so that original equipment manufacturers (OEMs) could spin different products quickly.
- A DSP-based baseband approach can cope better with different radio frequency (RF) and mixed-signal offerings which occur due to technology improvements and market changes; for example, automatic gain control (AGC) and AFC will change with different front-ends.
- Spare DSP MIPS come for free and enable product differentiation (echo cancellation, speech recognition, noise cancellation, better equalizers).

## TRENDS IN LOW-POWER DSPS

The digital baseband section is critical to the success of wireless handsets and, as mentioned previously, programmable DSPs are essential to

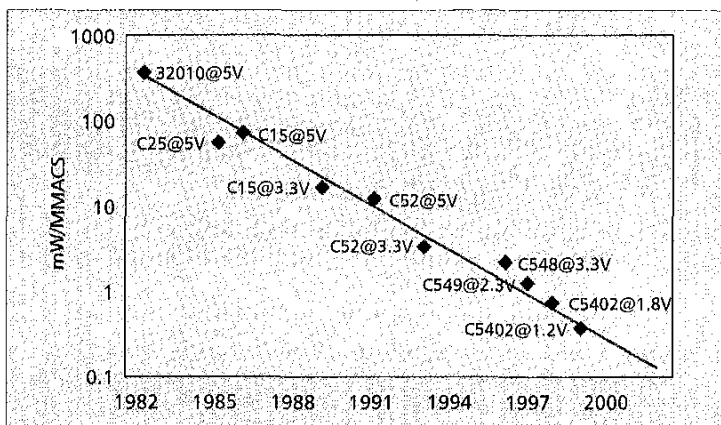


■ Figure 2. GSM function partitioning.



■ Figure 3. Layer 1 and application MIPS with time.

provide a cost-effective, flexible upgrade path for the variety of evolving standards. Architecture, design, and process enhancements are producing new generations of processors that provide high performance while maintaining the low power dissipation necessary for battery-powered applications. Many communications algorithms are multiply-accumulate (MuAcc) intensive. Therefore, we evaluate DSP power dissipation using milliwatts per million MuAccs (MMuAcc), where a MuAcc consists of fetching two operands from memory, performing a multiply-accumulate, and storing the result back in memory. As shown in Fig. 4, DSP power dissipation is following a trend of halving the power every 18 months [3]. As the industry shifts from second-generation to 3G wireless we see the percentage of the physical layer MIPS that reside in the DSP going from essentially 100 percent in today's technology for GSM to about 10 percent for wideband code-division multiple access (WCDMA). However, the trend shown in Fig. 4, along with more efficient architectures and enhanced instruction sets, implies that the DSP of three years from now will be able to imple-



■ Figure 4. Power dissipation trends in DSP.

ment a full WCDMA physical layer with about the same power consumption as today's GSM phones.

Since these DSPs use static logic, the main power consumption is charging and discharging load capacitors on the device when the device is clocked. This dynamic (or switching) power dissipation is given by:

$$\text{Power} = \alpha C V_{\text{swing}} V_{\text{supply}} * f$$

where  $\alpha$  is the number of times an internal node cycles each clock cycle, and  $V_{\text{swing}}$  is usually equal to  $V_{\text{supply}}$ . The dynamic power for the whole chip is the sum of this power over all the nodes in the circuit. Since this power is proportional to the voltage squared, decreasing the supply voltage has the most significant impact on power. For example, lowering the voltage from 3.3 V to 1.8 V decreases the power dissipation by a factor of 3.4. However, if the technology is constant, lowering the supply voltage also decreases performance. Therefore, technology scaling (which decreases capacitance) and power supply scaling are combined to improve performance while decreasing the total power consumption of the DSP, as shown in Fig. 5 [3]. The current TMS320C54x is capable of 120 MMuAccs at 1.8 V and 0.72 mW/MMuAcc in 0.18  $\mu\text{m}$  CMOS technology.

In what follows we use the TI C54x [4] as an example of an evolving DSP that is optimized for wireless applications. However, the reader should note that because of the growing importance of the wireless market (more than 400 million units projected for 2000 [5]), there are now several DSPs on the market that have been designed with wireless applications in mind, for instance, the Lucent 16000 series [6] and the ADI21xx series. This level of effort by several companies is a sign that the collective wisdom of the marketplace has chosen to bet on a programmable DSP future for wireless technology. We should also note that, although designed for wireless applications, these DSPs are finding major markets in other low-power applications such as telephony modems, digital still cameras, and solid-state audio players.

As mentioned earlier, the power difference between DSP and ASIC solutions was significantly reduced by designing the DSP for low-power wireless applications. Several power saving features are built into the C54x architecture and instruction set to reduce the code size and processor cycles required. The core uses a modified Harvard architecture that incorporates three data memory buses (two read, one write) and one program memory bus coupled with two data address generators and one program address generator. This architecture leads to high memory bandwidth and enables multiple operand operations, resulting in fewer cycles to complete the same function.

Another strategy used by DSP designers is to add instructions that, although fairly generic in themselves, allow efficient implementation of algorithms important to wireless applications. For instance, in the C54x, one of the arithmetic logic unit (ALU) inputs can be taken from a 40-bit barrel shifter, allowing the processor to perform numerical scaling, bit extraction, extended arithmetic, and overflow prevention. The shifter

and exponent detector enable single-cycle normalization of values and exponential encoding to support floating-point arithmetic for voice coding. A compare-select-store unit contains an accelerator that, for channel decoding, reduces the Viterbi "butterfly update" to four cycles. This unit generally provides acceleration for any convolutional code based on a single shift register, which accounts for all the codes commonly in use in wireless applications today.

The C54x instruction set also contains several dedicated instructions, including single and block repeat, block memory move, conditional instructions, Euclidean distance calculation, finite impulse response (FIR), and least mean square (LMS) filtering operations.

All these customizations permit power efficient implementations of 7.4 mW for IS-54/136 VSELP and 1.3 mW for the GSM speech coder. The trend toward more specialized instructions will continue increasing as the cost of supporting these instructions goes down. Other instruction enhancements for bit manipulation, which is traditionally done much more efficiently in ASIC, will occur in the near future.

Another trend in DSP evolution is toward VLIW processors to support a compiler-based programmer-friendly environment. Examples of this include TI's TMS320C6x [7], ADI's TigerSHARC [8], and Lucent and Motorola's StarCore [9]. These VLIW processors use explicitly parallel instruction computing (EPIC) with predication and speculation to aid the compilers. The processors are also statically scheduled, multiple-issue implementations to exploit the instruction-level parallelism inherent in many DSP applications. Although the application of this to physical layer processing in the handset is not apparent so far, these devices allow very efficient compilation of higher-level code, thus reducing the need for DSP-specific assembly-level coding of algorithms. As explained later, the trend of wireless toward an open applications-driven system will make this kind of DSP much more compelling as a multimedia processor in the handset.

Power management is very important in a low-power DSP, and the C54x utilizes a hybrid power management strategy with automatic local clock gating and three user-controlled idle modes to shut down the CPU only, CPU and peripherals, or a complete device shutdown preserving the memory states. A flexible digital phase-locked loop (PLL)-based clock generator and multiplier allows the user to optimize the frequency and power for their application. In general, these techniques allow a DSP not designed for a specific function to optimize its power usage for that function, bringing its power level closer to that of a dedicated ASIC design.

## COPROCESSORS

In this section we discuss how coprocessors can complement the function of programmable DSPs in the implementation of a flexible 3G platform. We will use the ARIB/European Telecommunications Standards Institute (ETSI) WCDMA standard [10] to illustrate our arguments. However, the points we make are quite general for

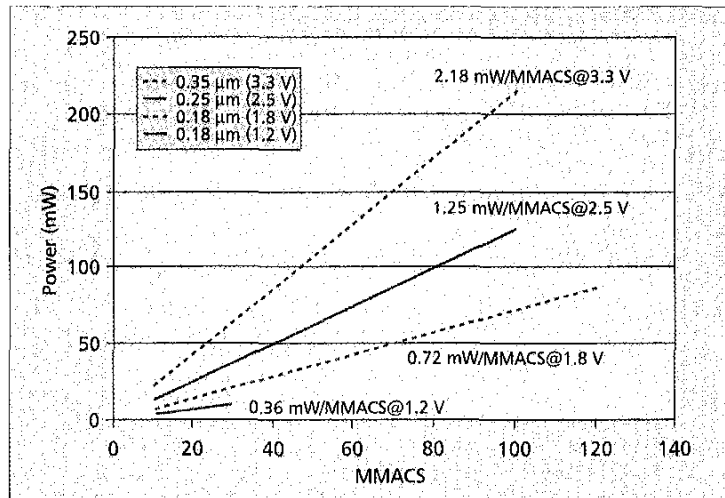
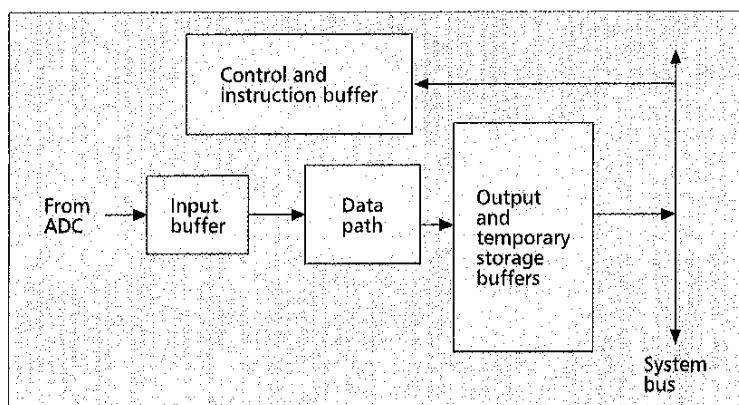


Figure 5. C54x MMACS vs. power

communications physical layer design. Although the second section implies that, for a given wireless standard, the amount of the physical layer performed by a DSP will increase to close to 100 percent as the standard matures, we are still faced with the problem of implementing a new standard with today's DSPs. Generally, because standards are driven by what is possible for ASIC implementation at a given power and cost point, a newly defined standard cannot be implemented in a DSP alone. For a WCDMA voice rate terminal, if we make a rough count of the "operations" required, only 10 percent are suitable for implementation on a current DSP, these being essentially the functions operating on data at the symbol rate as opposed to the chip rate. But a fixed function solution would be a high-risk option due to a lack of flexibility, especially in a new standard. Therefore, the system designer is faced with the problem of balancing the power and flexibility requirements. If we assume a long-term trend to increased use of more powerful DSPs, the designer also requires a roadmap for his/her design to migrate toward these devices.

One appealing solution to this problem is a coprocessor-based architecture with a single programmable device at its core. An example of such an approach is the Pleades project [11]. They have demonstrated a RISC engine with an attached configurable array of multipliers which is capable of doing a 16-point complex radix 2 fast Fourier transform (FFT) with a delay energy product for the function that is 0.02 percent and 4 percent of the same function on a Strong ARM and TMS320C54x, respectively. In a complete system the gains may not be as significant, but this nevertheless shows that there is significant potential in this approach. In this case the coprocessor provided significant performance gain for any algorithm that could be mapped to its multiplier array architecture. The gain is therefore only apparent for a domain-specific set of algorithms. But within this set it allows a good degree of flexibility.

We divide the world of coprocessors into loosely coupled and tightly coupled, which are defined relative to the average time to complete



■ **Figure 6.** A correlator coprocessor for chip rate processing in WCDMA.

an instruction on the DSP. With a tightly coupled coprocessor (TCC) the DSP will initiate a task on the coprocessor that completes in a few instruction cycles. A task initiated on a loosely coupled coprocessor (LCC) will run for many instruction cycles before it requires more interaction with the DSP.

The TCC will generally have a specific interface to the DSP core and will have access to some registers within that core. Since each task only takes a few cycles, it will naturally only involve a small amount of data. Also, parallel scheduling of tasks on the DSP and TCC will be very difficult, since the DSP will interrupt its task after a few cycles to service the TCC. Therefore, the DSP will generally freeze during the operation of the TCC. The TCC is therefore a user-definable instruction set enhancement that provides power and speed improvements for small tasks where there is no data bottleneck through the DSP. A TCC also may have a very specific task and be relatively small compared to the DSP. With time, the function of the TCC may be absorbed into the DSP by either replacing it with code in a faster, lower-power DSP, or absorbing the function of the TCC into the core of the DSP and giving it a specific instruction. An example of this sort of function would be a Galois arithmetic unit for coding purposes or a bit manipulation coprocessor providing data to symbol mappings that are not presently efficiently implemented in the DSP instruction set.

LCCs are more analogous to a subroutine call than an instruction. As they perform many operations without further DSP intervention, they will generally operate on large data sets. Unlike the TCC, the LCC will have to run in parallel with the DSP if it is to achieve its full benefit. This means that the programmer will have to be more careful with the scheduling of LCC instructions. However, since the LCC has minimal contact with the DSP, this should not be a problem. The main advantage of the LCC is that it solves the serious problem of bus bandwidth that can occur when either the raw input data rate to the system is very high, or else the number of times data is reused in calculations is very high. In either case, the bus bandwidth becomes the bottleneck to performing the computation because the data is stored

at the other end of the bus from the computational units. An LCC removes this bottleneck by having the computational units local to the data and arranged specifically for the data access required for a class of computations. In time the DSP will evolve to a point where its bus bandwidth and computational power are sufficient for the LCC's task and the pseudo subroutine implemented by the LCC will become a real subroutine.

The LCC concept applies easily at the chip rate to symbol rate boundary of a CDMA system. In the WCDMA physical layer the DSP would still perform many of the symbol rate processing tasks such as the timing recovery, frequency and channel estimation, and finger allocation. The chip rate processing tasks such as despreading, path delay estimation, and acquisition would be farmed out to a coprocessor that is designed to perform such tasks efficiently. One possibility is a correlator coprocessor, which performs the common despreading tasks for fingers and path delay estimation. The coprocessor can also perform some simple but high-MIPS tasks that occur directly at the chip to symbol boundary. Examples of these are coherent and noncoherent averaging for channel estimation. However, the DSP still chooses the type of averaging that should occur and how to post process the data to produce the final channel estimate. In effect, the system is fully programmable within the domain of CDMA chip rate processing. The DSP also has control of how the despreading MIPS are allocated. For instance, the DSP may choose to allocate a portion of the MIPS to one user with 6-multipath. Alternatively, it may reallocate these same MIPS to several users with less multipath to despread. Apart from allowing different users to differentiate and improve their WCDMA solutions completely in software, this allows the same system to be reprogrammed to perform WCDMA, GSM, CDMA2000, IS95, Global Positioning System (GPS), and other CDMA-based demodulation systems. It also provides a common platform for both low-cost voice-only terminals and high-end multimedia terminals. A simplified block diagram of this coprocessor is shown in Fig. 6. Note that the coprocessor is connected directly to the front-end to remove the chip rate data completely from the bus. Another important feature is that the instruction and output buffers are memory mapped to allow flexible access to the coprocessor by the DSP. Comparisons to fixed function designs show that, with careful design of the coprocessor logic, there is no significant power penalty to be paid for the flexibility. This is essentially because the data flow dominates the power budget, and this is independent of the flexibility of the design. The size of the coprocessor is somewhat larger than a dedicated design, but not significantly so within the complete system budget.

Decoding is another area which can benefit from the application of LCCs. Voice rate Viterbi decoding is easily performed on today's DSPs, but the higher data rate requirements in 3G make decoding hard to do programmably. Nevertheless, it is possible to find a DSP/coprocessor partition that maintains the flexibility required

along with a reasonable MIPS level on the DSP. As an example, for Viterbi decoding in the base station, the DSP could perform all the data processing up to the branch metric generation, and a coprocessor could perform the remaining high-MIPS tasks of state metric update and trace-back. This allows the DSP to programmably define a decoder for any code based on a single shift register, including puncturing to other rates. We estimate that to support 240 3G voice channels (with constraint length  $K = 9$ ) would require a 45K gate coprocessor and only 6 MHz of DSP processing on a C6x.

## APPLICATIONS AND ARCHITECTURES FOR FUTURE WIRELESS DEVICES

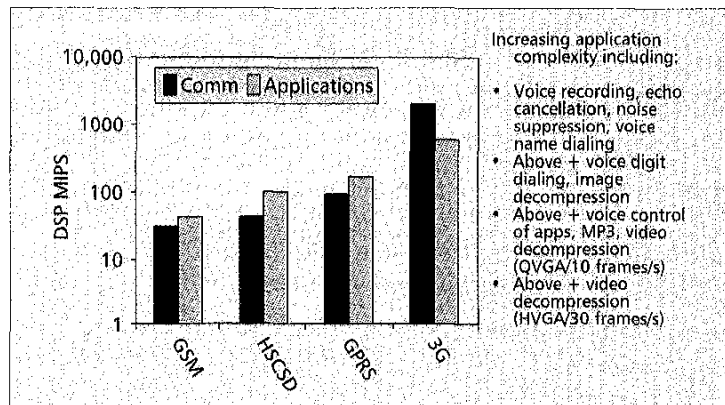
There have been several trends which have dominated cellular communications since its commercial debut in 1983. First is the trend from analog to digital. This has been a boon for cellular carriers because it has allowed them to triple their subscriber capacity with little to no increase in cost. However, the biggest impact of digital technology is yet to come. The transition to digital technology is also driving a transition from voice-only cell phones to an environment which can deliver a variety of data services to a broad range of mobile devices. Bits, after all, are just bits, and they can be used to transmit data as well as voice.

The key to this transition is the fact that the evolution of cellular technology from first to second and soon to third generation technology is mirroring the evolution of wireline systems by delivering more and more bandwidth to mobile subscribers. This bandwidth lags wireline technology by several years, but the year 2000 should represent a watershed for wireless technology as available bandwidths exceed 100 kb/s and packet-switched networking technology is introduced. Although voice communications will probably always be the dominant mobile service, the next century will begin to see a proliferation of targeted data services for mobile subscribers.

Examples of such services could include:

- Imaging services: Such capabilities would allow a businessperson to transmit the image of a document he/she is discussing. It would allow a real estate agent to show a client the image of a new house or a grandmother to receive the latest snapshots of a new baby via e-mail.
- Location-based services would allow travelers to locate and navigate to new destinations. It would allow advertising to be delivered to mobile subscribers in a specific area. It would even allow parents to keep track of their children.
- Audio and visual entertainment could be delivered wirelessly to mobile subscribers. You could even imagine mobile subscriber-to-subscriber gaming environments.

The best of these applications are yet to be conceived. Those applications will come when low-cost high-bandwidth packet networks are generally available. A common theme in all of these services will be an insatiable thirst for sig-



■ Figure 7. Applications drive DSP MIPS.

nal processing power. Figure 7 illustrates that even as communication processing needs become more demanding, applications requirements will also drive a need for more powerful DSPs.

Traditional cell phones use a dual processor architecture to implement what is in effect a simple modem. Let's call that the communications engine. Data-centric terminals will have new challenges and potentially new architecture. Top-of-the line phones developed in 1999 are starting to integrate personal organizer and information exchange features [12, 13], and this will lead to wireless phones with open operating systems onto which a user can download third-party applications. Mobile devices which deliver data services will have to add processing resources to support an increasingly complex user interface, more sophisticated operating environments, and support for added applications beyond voice. Let's call those resources the applications engine.

Starting with a phone that implements the communications engine using a dual-core RISC + DSP architecture, there are several options for adding resources to support data services. Most current approaches incorporate a single RISC core in addition to the communications engine. Other options include multiple DSPs or even hardware coprocessors to accelerate the heavy additional media processing required to implement the applications engine.

In summary, support for future data-centric mobile devices will require a new approach to the architectures that implement these 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 RISCs, which implement protocol stacks, user interfaces, and high-end operating systems. You will also see programmable DSPs and coprocessors, which can provide power-efficient media processing, and support needed application flexibility and upgradeability. The biggest open issue is the extent to which the applications and communication functions can be effectively combined in one programming environment.

*Support  
for future  
data-centric  
mobile devices  
will require a new  
approach to the  
architectures  
that implement  
these devices.*

## REFERENCES

- [1] <http://www.mot.com/SPS/WIRELESS/products/DSP56652.html>
- [2] <http://www.ti.com/sc/docs/wireless/97/digbase.htm>
- [3] P. Gelabert and T. Stetzler, "Industry's Lowest Power General Purpose DSP," Embedded Processor Forum, May 3-4, 1999.
- [4] <http://www.ti.com/sc/docs/products/dsp/c5000/index.htm>
- [5] Dataquest "Mobile Communications Semiconductor Applications Markets, 1997-2002," Apr. 12, 1999, WSAM-WW-MT-9901.
- [6] <http://www.lucent.com/micro/wireless.html>
- [7] <http://www.ti.com/sc/docs/products/dsp/c6000/index.htm>
- [8] <http://www.analog.com/industry/dsp>
- [9] <http://www.starcore-dsp.com>
- [10] <http://ftp.3gpp.org>
- [11] <http://infopad.EECS.Berkeley.EDU/research/reconfigurable>
- [12] <http://www.qualcomm.com/pdq>
- [13] <http://www.symbian.com>

## BIOGRAPHIES

ALAN GATHERER (gatherer@ti.com) received his B.Eng. degree in electronic and microprocessor engineering from Strathclyde University, Scotland, in 1988, and his MS and Ph.D. degrees in electrical engineering from Stanford University in 1989 and 1993. Since 1993 he has worked at Texas Instruments. Presently he is a senior member of technical staff and manager of the wireless communications group within corporate research. He is involved in the development of WCDMA technology for 3G cellular tele-

phony. He holds three patents in the area of digital communications.

TRUDY STETZLER received her B.S. in electrical engineering from Pennsylvania State University in 1984, her M.S. in electrical engineering from the University of California, Berkeley, in 1985, and her M.B.A. from the Wharton School, University of Pennsylvania, in 1997. In 1997 she joined Texas Instruments' DSP New Business Development Group as a senior member of technical staff performing system-level analysis. Prior to 1997, she was a Distinguished Member of Technical Staff at Bell Laboratories designing analog and RF ICs for cellular phones.

MICHAEL L. MCMAHAN received an S.B. in electrical engineering from the Massachusetts Institute of Technology, and an M.S. in computer science from the University of Illinois. Since joining TI he has worked in the areas of speech and digital signal processing including the TI Speech product for the TI and IBM PC's and Project Watson, which developed speaker verification and recognition technology for Sprint's long distance telephone network. Currently he is a TI Fellow, responsible for worldwide R&D for the Wireless Business Unit of TI.

EDGAR AUSLANDER (E-Auslander@ti.com) is worldwide strategic marketing director at Texas Instruments' Wireless Communications Business Unit, Dallas, Texas. He obtained his M.E.Eng. and M.B.A. degrees from Cornell University and Columbia Business School in 1988 and 1990, respectively. He first worked at Texas Instruments in Nice, France, as European product marketing manager for the TMS320 Digital Signal Processors product line. He became TI's worldwide GSM marketing manager in 1993 prior to his current position.