Software-Defined Radio Prospects for Multistandard Mobile Phones

Ulrich Ramacher
Infineon Technologies AG

A multitude of existing and evolving cell phone standards, coupled with the need to react quickly to market requirements, is radically changing the system architect's task. Two alternative approaches to classic ASIC-centered and DSP-assisted designs have emerged: reconfigurable architectures and DSP-centered and accelerator-assisted architectures.

Multiple standards have become the norm in the high-end mobile phone market. Cellular networking standards such as the Global System for Mobile Communications (GSM), Wideband Code Division Multiple Access (WCDMA), and High-Speed Downlink Packet Access (HSDPA), wireless Internet standards such as 802.11a/b/g, and mobile TV standards such as Digital Video Broadcasting-Handheld (DVB-H) are among many now in use, while Bluetooth technology supports local connections to headsets and other equipment.

Next-generation mobile phone standards are on their way. By 2010 to 2012, high-end cell phones must be able to meet the additional demands of Long-Term Evolution (LTE), Worldwide Interoperability for Microwave Access (WiMax), and ultra-wideband (UWB).

This multitude of existing and evolving standards, outlined in Table 1, coupled with the need to react quickly to market requirements, is radically changing the system architect's task.

BASEBAND SOLUTIONS

In the recent past, silicon area and power consumption were the sole design criteria. On the baseband side, this led to architectures in which application-specific circuit blocks, or macros, dominated the computation-intensive signal-processing parts and digital signal processors (DSPs) managed the layer 1 control part, as the bottom row of Figure 1 shows. However, with the advent of new standards and the shift to ubiquitous communication, continuing this style of design results in an intolerable increase in the number of macros required.

Two alternative approaches to classic ASIC-centered and DSP-assisted architectures have emerged: reconfigurable architectures and DSP-centered and accelerator-assisted architectures.

Reconfigurable architectures

This approach, shown in the middle row of Figure 1, involves

- identifying those algorithmic functions, and even pieces thereof (subfunctions), that a set of standards share, and
- using reconfigurable data paths to emulate these common subfunctions.

Since use cases for high-end and mid-level mobile phones involve only a few simultaneous standards, it is believed that reconfigurable architectures would achieve sufficient area savings, with only slightly increased power consumption because of the inherent multiplexing of subfunctions. This technology also appeals to traditional baseband architects because it deviates least from what they are accustomed to. Numerous startup companies accordingly adopted this approach in the mid-1990s, but

Published by the IEEE Computer Society
0018-9162/07/$25.00 © 2007 IEEE

Table 1. Mobile phone standards.

Specification	Wi-Fi ultra-wideband	802.11a/b/g	802.11n	Wireless broadband (WiBro)	Mobile WiMax (802.16—2005)	3G LTE (cellular WAN)	Digital Video Broadcasting—Handheld	Digital Video Broadcasting—Terrestrial
Application	High-speed local interconnect, wireless USB	Medium-speed LAN	High-speed LAN	Mobile wireless access	Mobile wireless access	Mobile data/voice	Mobile TV	Mobile TV
Range	10 m	80 m	50-150 m	1-5 km	1-5 km	1+ km	Broadcast	Broadcast
Rate	480 Mbps	11 Mbps (b), 54 Mbps (a/g)	100-600 Mbps	3-50 Mbps (downlink)	63 Mbps (downlink)	100 Mbps (downlink)	384 Kbps	7 Mbps
Frequency	3.1-10.6 GHz	2.45/5.8 GHz	2.45/5.8 GHz	2-6/2, 3 GHz	2-6/2, 3 GHz	1.25/2.2/5/10/20 GHz	0.8 MHz, 1.6 GHz	0.8 MHz, 1.6 GHz
Modulation	Orthogonal frequency division multiplexing	Direct-sequence spread spectrum/ complementary code keying, carrier sense multiple access, orthogonal frequency division multiplexing	Carrier sense multiple access, orthogonal frequency division multiplexing	Orthogonal frequency division multiple access	Orthogonal frequency division multiple access	Orthogonal frequency division multiple access/single carrier frequency division multiple access	Orthogonal frequency division multiple multiplexing	Orthogonal frequency division multiple multiplexing

Table adapted from a figure provided by NXP Semiconductors.

a success story is still lacking for mobile phones.

Reconfigurable architectures exhibit a degree of flexibility that depends on the chosen granularity and number of common functions and subfunctions, respectively. The most fine-grained set is obtained by decomposing algorithms into their basic subfunctions: add, subtract, multiply, divide, and so on. These elements must then be diversely concatenated to produce the elementary algorithms, such as a fast Fourier transform (FFT), of the application under consideration—the best implementation being a classic general-purpose register-based DSP architecture. However, even with multiple DSPs to reduce power consumption, this extreme still consumes too much power and is thus intractable for handhelds.

The coarsest-grained set of functions, consisting of the elementary communication

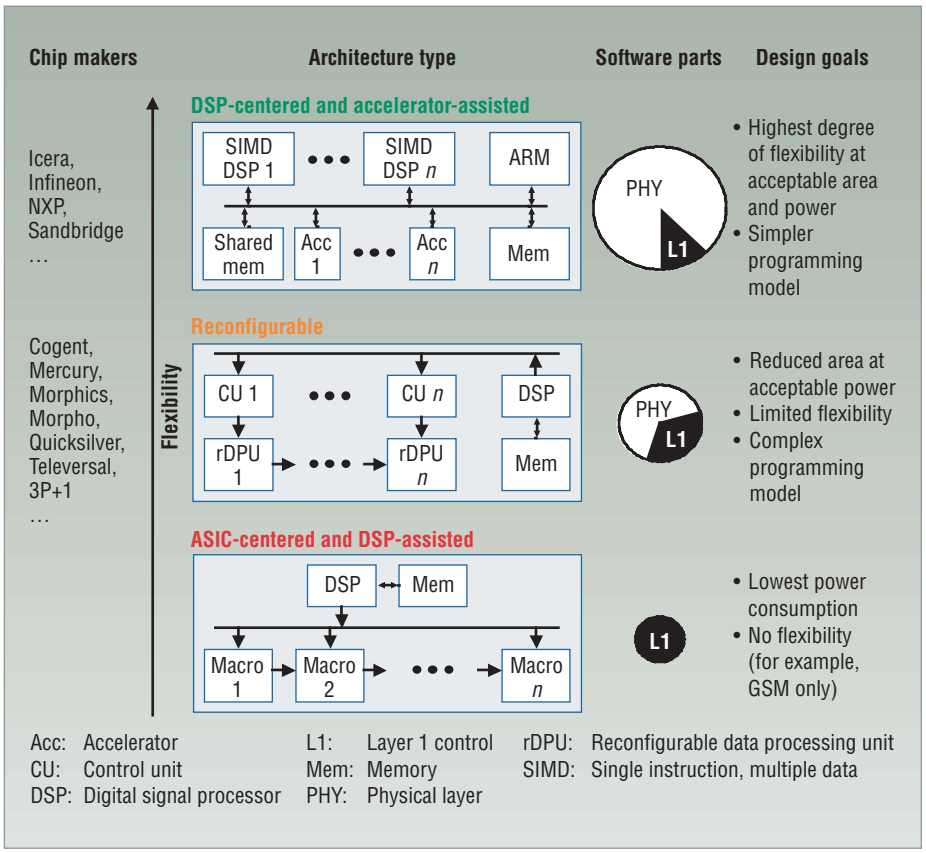

Figure 1. Baseband architectures for multistandard mobile phones. Two alternative approaches to classic ASIC-centered and DSP-assisted architectures have emerged: reconfigurable architectures and DSP-centered and accelerator-assisted architectures.

algorithms in the selected set of standards, represents the other extreme. However, since diverse standards share few elementary algorithms, implementation is impractical because of the large silicon area the many different macros require.

Between these two extremes lie reconfigurable data path structures. As these data paths are not simply attached to registers, but to switch-controlled buffers or latches that connect different data paths directly, there is no easy way to program them; indeed, this has been a major problem.[1]

A further obstacle is the lack of specification of the degree of flexibility needed—neither the customer nor marketing analysts provide the system designer with a clear answer. On the contrary, the ideal is a flexible architecture that "proactively" anticipates the needs for future change.

An SDR baseband solution will always consume more power than a classical one.

DSP-centered and accelerator-assisted architectures

Few engineers have attempted to design a truly programmable DSP architecture. For power consumption reasons, it must comprise multiple DSPs running at a few hundred MHz; for area, power, and performance reasons, it must use lean general-purpose DSPs with special instructions assisted by a small number of accelerators (macros or application-specific instruction processors). Such an architecture, shown in the top row of Figure 1, is the converse of the ASIC-centered and DSP-assisted architecture.

A DSP-centered and accelerator-assisted architecture offers the highest possible degree of flexibility, but power consumption is probably too high for use in cell phones. However, many software-defined radio proponents claim SDR cell phones would reach competitive power figures in 90 nanometers,[2,3] and there is widespread belief that programmable multiprocessor solutions will replace today's ASIC-centered and DSP-assisted phones, the only question being when.

RADIO-FREQUENCY SOLUTIONS

The RF side of the problem of multistandard mobile phones is similar to that of the baseband side. A transceiver solution is typically developed for each standard and often comprises separate paths—low noise amplifier (LNA), mixer, and so on—for each band, on the same RF integrated circuit (RFIC). In the high-end market segment, individually optimized cellular, wireless local area network (WLAN), TV, and Bluetooth chip sets are mounted on the motherboard, together with switches, filters, power amplifiers (PAs), and LNAs as well as their baseband companions. As more standards are definitely to come, a better solution is needed.

Compared to the baseband side, there are few proposals for an SDR front end,[4,5] and fully developed SDR concepts, let alone chips, are not yet available. In the meantime, a more moderate approach based on a configurable analog-to-digital converter, filters, a digital interface, a common voltage-controlled oscillator and synthesizer, and perhaps a common mixer seems to be amenable for SDR RF products beginning in 2010. Accordingly, realizing SDR cell phones boils down to implementing the use cases for multistandard mobile phones with the minimal number of wideband transceivers and multistandard baseband chips.

As the SDR RF solution is still under research, the focus here is on the SDR baseband solution.

SDR TECHNOLOGIES AND MARKET POTENTIAL

Philips Semiconductors, now NXP Semiconductors, acquired Systemonic, the first start-up to propose SDR processor solutions, in 2002. NXP's second-generation embedded vector processor (EVP)[6] became available in silicon in January 2007 and is planned for use in connectivity and cellular products from 2008 onward. Like Systemonic's On-DSP, EVP's architecture belongs to the single-instruction, multiple-data (SIMD) class. Two other start-ups with SDR baseband offerings, Sandbridge and Icera, were founded in 2001 and 2002, respectively. Both Sandbridge's Sandblaster[1] and Icera's Livanto[3] are also in the SIMD class.

Area consumption

Although some figures for area and power consumption are available, it is not possible to perform a fair comparison of the various solutions unless a benchmark for multistandard operation of cell phones is defined. Until then, product announcements are the best grade of measuring the maturity of SDR baseband solutions. Icera and Sandbridge seem to be involved in designing PC cards and PDA cell phones, but not mobile phones, suggesting that known offerings still consume too much power. Of course, an SDR baseband solution will always consume more power than a classical one. However, 50 to 100 milliwatts higher power consumption, depending on the use cases, might be affordable and be outweighed by area savings.

The Nokia 6633 cell phone provides a gauge point for area comparison of SDR with classical solutions. This device currently uses a 51mm^2 baseband chip in 90-nm complementary metal-oxide semiconductor (CMOS) process technology for GSM/Enhanced Data Rates for GSM Evolution (EDGE) and WCDMA. At least 6 mm^2 more would be needed to support HSDPA, another

8mm² for an 802.11a/b/g baseband chip, about 10 mm² more for a DVB-H baseband chip, and about 16 mm² more for a WiMAX baseband chip. The green line in Figure 2 represents this multiple-radio configuration using 65-nm technology. Integrating all the standards on a single chip would consume even more area.

The black line gives the area for the MuSIC (multiple SIMD core) chip, Infineon's SDR prototype solution, for the various use cases. MuSIC contains all the elements of a baseband solution: a digital interface to the RF front end, a set of SIMD DSPs and accelerators, shared memory, a bus bridge, an Advanced RISC Machine (ARM) processor subsystem (for protocol stack processing and execution of the multiprocessor OS) with peripherals, and an interface to the multimedia processing site.

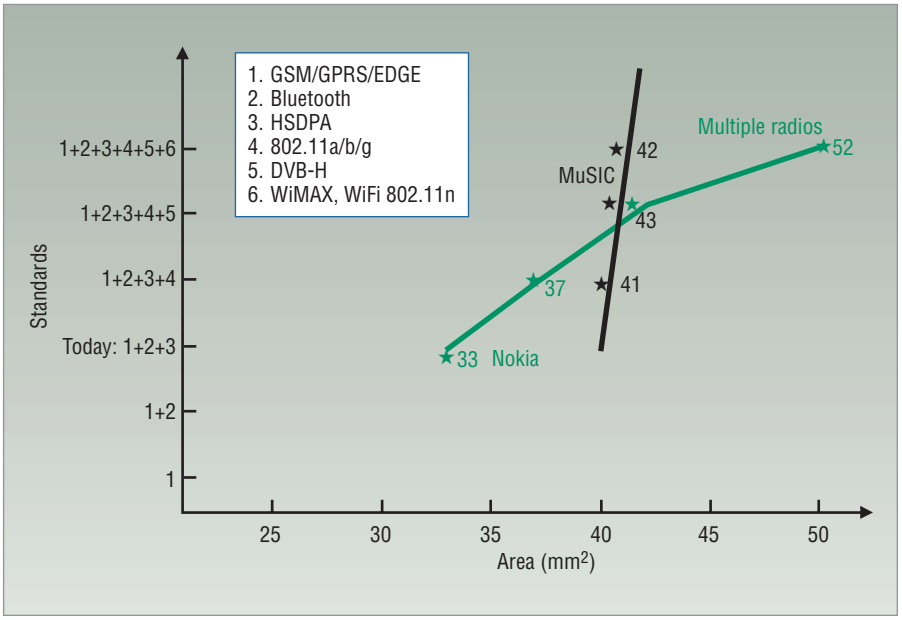

Figure 2. Area consumption of multiple radios versus SDR radios in 65-nm CMOS process technology.

Consider the use case with Bluetooth and GSM/Universal Mobile Telecommunications System (UMTS) monitoring in parallel and with a cellular, WLAN, or one of the many mobile TV standards being active. SDR breaks even in area terms with a multiple-radio solution for cellular networking, wireless Internet access, and mobile TV access, and it is definitely superior when also considering WiMax, LTE, and UWB.

Scalability

It is apparent that the SDR baseband solution does not scale well below the break-even point. Luckily, these market segments comprise today's entry-level and ultra-low-cost phones, which over the next few years will incorporate the feature set formerly reserved for the mid-level market. SDR might therefore still be worthwhile because the area difference between 33 and 42 mm² of the break-even SDR solution is not prohibitive compared to the enormous increase in benefit for the end user: Internet and TV access.

In the other direction, SDR is definitely scalable to the computational needs of the looming LTE, WiMax, and UWB standards, even though particular accelerators for multiple-input-multiple-output processing (MIMO) or low-density parity-check (LDPC) codes might be necessary. The key metrics are scalability by frequency, number of processors, and software, respectively.

Because of their unprecedented flexibility and scalability, SDR baseband solutions for all regions and market segments except the low end can be obtained from a single-system platform comprising hardware and firmware. Country-specific selection of standards, either preconfigured or downloadable, are manageable. Even a software update to recently developed standards will be possible.

By 2010, around 70 million phones in the high-end segment and 240 million in the mid-level segment—a quarter of the total number of cell phones produced—will be required to offer multistandard capabilities. These market segments would benefit considerably from an SDR baseband solution provided power consumption is also competitive. Based on measurements of Infineon's MuSIC-1 chip obtained for peak loads like FFT, finite impulse response (FIR), and complete standards like 802.11b as well as WCDMA release 99, it can be stated that SDR baseband solutions for mobile phones are clearly competitive with respect to power consumption and area in 65-nm CMOS.

MUSIC-1: A PROTOTYPE SDR BASEBAND SOLUTION

Figure 3a shows the block diagram of a multistandard radio system featuring two RF chips for the different standards and a single baseband processor comprising multiple DSPs and accelerators and an ARM processor for protocol stack processing. The baseband DSPs, highlighted in Figure 3b, were designed as SIMD cores. This kind of DSP core is particularly suited for the computationally complex algorithms in communication systems, which exhibit substantial data parallelism.

The SIMD core, shown in Figure 3c, is based on previous work[1] and has been simplified as well as extended for use in communication systems. The cluster of SIMD cores is accompanied by dedicated programmable processors

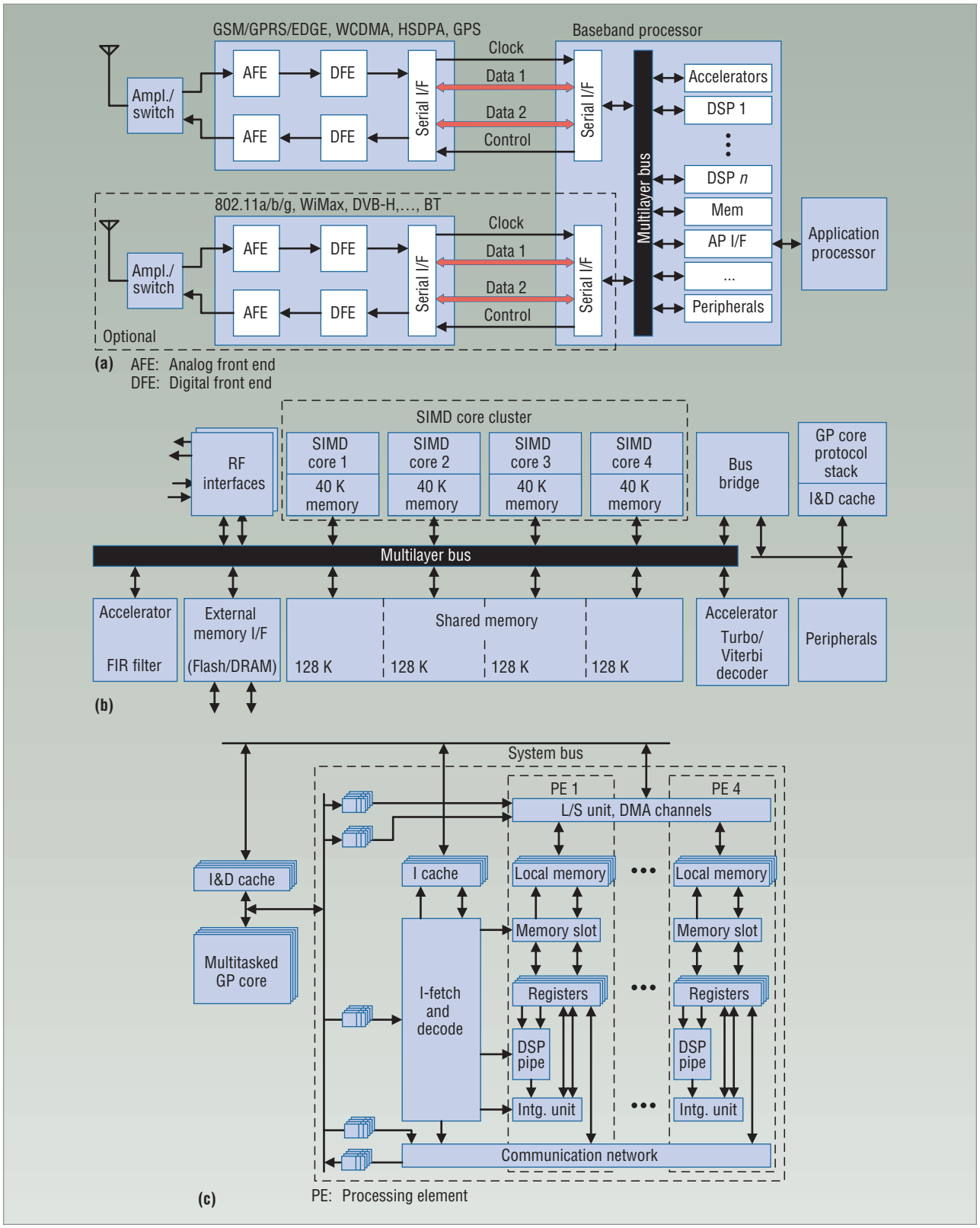

Figure 3. Infineon's MuSIC-1 chip. (a) Block diagram of SDR modem. (b) Baseband DSP. (c) SIMD core.

for filtering operations as well as channel encoding and decoding. A general-purpose processor along with a set of standard peripherals executes the L1 control and protocol stack software. External memories provide the code and data storage. A couple of RF interfaces connect the baseband processor to the digital front ends and provide

data buffers that adapt the baseband software's block-oriented operation mode to the front ends' streaming mode. The interfaces are highly configurable with respect to block sizes and transfer modes.

All of the baseband processor's components share an on-chip memory, which consists of multiple physical banks to support simultaneous accesses from different bus masters. Similarly, the system bus is organized in several independent layers for multiple simultaneous transfers.

Table 2. MuSIC power consumption, 300 MHz at 0.9 volts nominal.

Component/process	Power consumption, 6 base stations and 12 rake fingers (milliwatts)	Power consumption, 3 base stations and 8 rake fingers (milliwatts)
SIMD cores	239	159
Finite impulse response	33	33
TV access	27	27
ARM subsystem	24	24
Buses, memory, and synchronization	59	37
Total	**382**	**280**

The accelerators support various parameters that cover the requested communication standards. In addition, both accelerators observe two standards concurrently to meet the requirements as derived from the use case. Instead of implementing a separate macro for each of the supported modes, the application-specific instruction set processor (ASIP) accelerators are based on fine-grained instructions for arithmetic operations or bit manipulation. This retains the maximum level of flexibility for the entire platform and saves considerable power.

SIMD core

Each SIMD core contains four processing elements (PEs) and operates at a clock frequency of 300 MHz. The cores support special instructions like saturating operations and finite-field arithmetic as well as long-instruction word (LIW) features for performing arithmetic operations, accesses to the local memory, and data communication between the PEs in parallel.

The execution pipeline is four stages long, which makes it possible to implement complex instructions as well as relax timing requirements for the memory, reducing the memory's supply voltage and thereby the power consumption. However, a long pipeline suffers from stall cycles that are necessary to resolve data dependencies between instructions from the same task. Thus, the SIMD core is operated in hardware multitasking, where each of the four pipeline stages contains an instruction for one of four separate tasks. All storage elements like local memory and register file are replicated accordingly.

The local memory is fully visible to the programmer, in contrast to a conventional cache. The programmer can use separate load/store units, which provide several direct memory access (DMA) channels, to fully control transfers between the local memory and the shared on-chip memory outside the SIMD core. This helps in overlapping data transfers with computation, thereby increasing performance.

General-purpose core

Along with the PE array and its PE controller (for instruction fetch and decode logic), the system includes a scalar general-purpose core, which is multitasked in the same way as the PEs. It runs control programs providing the PE controller with addresses of instructions that reside in the shared on-chip memory. Likewise, the general-purpose core manages the load/store units by setting up memory addresses and so on for transferring data objects.

Controllers

Each of the controllers in the SIMD core, general-purpose core, PE controller, and load/store units accesses the shared memory through the system bus, which leads to stall cycles due to bus and memory conflicts. Misses in the instruction and data caches introduce further stall cycles. In a conventional synchronous arrangement, a stall condition in any of the controllers would halt the entire SIMD core.

To minimize stall performance penalties, the SIMD core uses first-in, first-out (FIFO) memories to decouple its controllers from one another. Consequently, rather than occurring in every clock cycle as in synchronous processors, synchronization between the controllers occurs explicitly at certain points that the programmer defines. Between such synchronization points, each controller can proceed at its own pace without being stalled by the others.[1]

Power consumption

MuSIC-1 was originally designed in 90-nm CMOS. Twenty-eight million logical transistors, 6 Mbits of SRAM, and six layers of wiring took 57 mm^2. On the basis of the power measurements and with the power shrink factor of 1.7 from 90- to 65-nm CMOS technology, the power figures shown in Table 2 were derived for the complete baseband solution, including accelerators, on-chip memory and buses, and an ARM 926 subsystem with peripherals. These figures hold for the complete WCDMA standard in real time in the worst case of having to cope with six base stations and 12 rake fingers, which requires eight SIMD cores. Normally, a cell phone communicates 60 percent of the time with one base station and 40 percent with two others. The third column in Table 2 shows the power figures obtained when three base stations and eight rake fingers are used, which needs only four SIMD cores.

These numbers are upper limits, as the MuSIC-1 prototype used regular V_{th}, six metal layers, and a single V_{dd} only. A further reduction of 20 to 30 percent seems feasible. Because MuSIC can implement WLAN, Bluetooth, and mobile TV (including H.264) standards as well, the maturity of SDR for use in 65-nm baseband solutions is evident.

R&D CHALLENGES IN SDR

Multiple processors for system-on-chip (MPSoC) technology is disruptive for cell phones because it replaces a traditional style of architecting baseband systems. Today's SDR baseband solutions are comparable to the first-generation CMOS photo sensors, which initially were regarded as being inferior to charge-coupled devices and, at best, suitable only for the low-performance market segment. Nowadays, leading camera manufacturers use CMOS photo sensors even in their high-end professional series. Similarly, SDR cell phones are expected to make an inroad in 2010 and to dominate from 2015 on. However, researchers must first overcome several R&D challenges that stem from the system's complexity.

Typically, 90 percent of what has been developed should be reused in the next architecture iteration.

Modeling

There has been little research to date on MPSoC-based mobile phones. This is not to downplay MuSIC's breakthrough achievement; it is to say that today's SDR baseband solutions are a first generation only. In the GSM era, system engineers mapped algorithms directly into hardware, but in SDR they must model algorithms and architectures separately to a large extent—for example, using Simulink and SystemC, respectively. The question of how fine-grained the emulation of standards—by reconfigurable data paths or by general-purpose DSPs with additional instructions and accelerators—should be depends on the flexibility needed for particular use cases and on the platform's scalability for various market segments, that is, its reuse.

Measuring the frequency of subfunctions at various granularity levels and determining existing data dependencies must occur independently of a hardware/software implementation and would suggest whether to choose a reconfigurable architecture or one of the DSP-centered architectures. However, the plethora of standards makes it difficult for one company to obtain this input; in the absence of such knowledge, the system architect often must rely on agreed-upon procedures for modeling algorithms and protocols, interchanging data, or generating C programs. For the time being, and given MuSIC's favorable power and area figures, the DSP-centered and accelerator-assisted type of architecture seems to offer the safer path to success.

Mapping

After opting for an entry point into the design space, the system architect must map subfunctions into hardware and software in a quantitative manner. Assuming a DSP-centered and accelerator-assisted type of architecture, a single, general-purpose DSP represents the most flexible choice. The limits of 90-nm CMOS dictate that any design should contain the smallest possible number of general-purpose DSPs working at the highest possible clock frequency and voltage to maintain the given power and area budgets. For example, taking 300 MHz as the clock frequency, at least 32 DSPs would be needed to meet WCDMA standards; as the resulting area would be too high, the design must add extensions to the general-purpose DSP instruction set or a set of accelerators. This approach ensures a simple programming model and highest flexibility while maintaining the area and power budget.

Optimization

A cycle-accurate simulator of the DSP core must be available from the beginning because partitioning the C code into tasks and threads is based on counting cycles. The goal is to maximize overall DSP throughput and to minimize the communication and data dependencies between DSPs. Optimization constitutes a high-dimensional problem, the solution for which results in the definition of the sequences of tasks and threads to be placed on each DSP, as well as the definition of the memory and bus architecture.

Design space exploration

In the course of iterative development of the SIMD architecture, the DSP simulator, compiler, and virtual prototype of the SDR baseband solution must be extended. However, companies too often lack the time and resources to step back on the chosen design trajectory. Typically, 90 percent of what has been developed should be reused in the next architecture iteration. Other issues hampering design space exploration of MPSoC concern bit-true virtual prototyping, which must be fast enough but not sacrifice cycle accurateness; design for minimal leakage, with idle and active power blending semi- and full-custom circuits at the transistor level in advanced CMOS technologies beyond 65 nm; and methodology and tools for code generation, design, verification, and test.

Budgetary constraints

Many researchers have addressed these technical issues, but often in isolation. For example, emerging networks-on-a-chip are scalable to the number of proces-

sors and memory, but NoCs are not recommended for cell phones because of their high power consumption and performance/latency disadvantages. Indeed, the bus and memory architecture seems to be the only piece which is not scalable in a multiprocessor architecture for multistandard cell phones.

Research on SDR cell phones must tackle the complete system—radio front end, baseband, and protocol processing. Integrating the results of research on parts of systems is insufficient. Such efforts might even be meaningless as MPSoC for mobile phones consists of strongly interacting components and thus requires a holistic approach to modeling, design, verification, and test.

Given that creating a next-generation architecture takes about four years, which is longer than a product architecture's lifetime, limited budgets as well as time-to-market pressures dictate that research and development of SDR cell phones occur simultaneously. In other words, the same team must be responsible for both invention and innovation. Cultivating a network of researchers and developers from industry as well as academia to collaborate over a long period of time is a prerequisite to success.

SDR baseband processors constitute a multiprocessor class of their own: similar in complexity to high-end microprocessors, and particular as to providing very high processing power of several giga-instructions per second for scalar operations and several tens of giga-operations per second for vector operations at very low power consumption (less than 0.5 watts).

For performance and power consumption reasons, an SDR baseband chip maker must design the L1 signal processing and control software. Consequently, a standardized hardware abstraction layer or virtual SDR machine is not needed—indeed, it would consume considerable power. The API between handset maker and chip manufacturer resides above the L1 software and firmware. This results in a greatly simplified transfer of L2/3 protocol software for service selection or handover scenarios to different SDR baseband chips (additional API standardization could help reduce it further), and lets handset makers avoid dependence on a single chip manufacturer.

SDR modem solutions represent a new stage in complexity. Semiconductor vendors, software houses, and major handset manufacturers must cooperate closely to support the various use cases, with the required levels of flexibility, while simultaneously achieving high enough volumes to justify the return on the huge investment needed. Given today's half-dozen main baseband chip manufacturers for cell phones and the high engineering costs for an SDR modem, three to four SDR modem platforms can be expected to survive through 2015.

Acknowledgments

The MuSIC team consists of Wolfgang Raab, Ulrich Hachmann, Dominik Langen, Hans Blüthgen, Cyprian Grassmann, Mirko Sauermann, and Matthias Richter (Siemens CT). Thanks to Bob Broderson (BWRC), Tobias Noll (TU Aachen), Günter Elst (FhG IIS-EAS), Rene Schüffny and Gerhard Fettweis (TU Dresden), and Hans-Gerd Thönnes (Infineon Design Team NRW), who performed pieces of MuSIC's design. Special thanks to my colleague Steve Hopper for his critical text reading. The project was funded in part by the German Federal Ministry of Education and Research (BMBF) under grant 01BU0650.

References

1. U. Ramacher et al., "A 53-GOPS Programmable Vision Processor for Processing, Coding-Decoding and Synthesizing of Images," *Proc. 27th European Solid-State Circuits Conf.*, Frontier Group, 2001, pp. 160-163.
2. J. Glossner, "The Sandbridge SB3011 Multiprocessor System on a Chip for Software-Defined Radio Handsets," presentation, 16 Aug. 2006, 6th Int'l Forum Application-Specific Multi-Processor SOC, Estes Park, Colorado.
3. S. Boland, "Architectures for Tomorrow's Mobile World," presentation, IET/FSA Int'l Semiconductor Executive Forum, 9 May 2006, Munich, Germany.
4. R. Bagheri et al., "An 800-MHz–6-GHz Software-Defined Wireless Receiver in 90-nm CMOS," *IEEE J. Solid-State Circuits*, Dec. 2006, pp. 2860-2876.
5. A. Kruth et al., "A Multimode Receiver Front-End for Software-Defined Radio," *Proc. 9th European Conf. Wireless Technology*, Horizon House Publications, 2006, pp. 19-22.
6. K. van Berkel et al., "Vector Processing as an Enabler for Software-Defined Radio in Handheld Devices," *EURASIP J. Applied Signal Processing*, issue 16, 2005, pp. 2613-2625.

Ulrich Ramacher *is responsible for SDR R&D at Infineon Technologies AG, Munich, Germany. His research interests include communication and multimedia systems, multiprocessor architectures, and system design methodology. Ramacher received a PhD in theoretical physics from the University of Tübingen, Germany. He is a member of the IEEE Computer Society. Contact him at ulrich.ramacher@infineon.com.*