The Modem Design Singularity

John Haine

22 December 2009

Cognovo Ltd
Melbourn Science Park
Melbourn, Royston
Herts. SG8 6HB
UK

T: +44 (0)1763 262477
sales@cognovo.com
www.cognovo.com

Summary

At the heart of a mobile phone is an Application Specific Integrated Circuit (ASIC) to do all the necessary complex signal processing, communications protocols, and applications. Mobile devices have always pushed the limits of what is possible with silicon technology. The next generations – LTE and LTE Advanced – will only be possible using the very latest VLSI processes with feature sizes of 45 or 32 nanometres.

A device for an LTE phone will roughly double the complexity of the generation that went before, measured by the number of computing operations required per second. At the same time, the costs of designing and fabricating a new chip using the latest processes are increasing exponentially. These two factors compound together to create the need for a new approach to designing the cellular modem – the core of the mobile device. We call this the “design singularity”.

Cellular modems have always tended to use conventional processor cores but to surround these with dedicated hardware for the very high speed signal processing needed, especially for broadband wireless systems such as HSPA+ and LTE. Dedicated hardware has been required because programmable processors just haven’t had the processing capability needed with reasonable power consumption. Using dedicated hardware makes development difficult and requires an expensive “re-spin” of the chip if errors are found. Moreover, it can be difficult to re-use the design in the next generation device because architectures have to change.

Now, a new type of computer architecture, the Vector Signal Processor, together with an advanced design tool-chain, makes it possible at last to match the performance of dedicated hardware with a software-programmable device, to create the “Software Defined Modem”. Using the VSP, designing the chip hardware can be separated from designing the modem; modem design errors can be corrected without hardware changes; and the software can be re-used for later device generations as the computing platform is scalable both in its basic architecture and how it makes use of advances in silicon process technology.

The Mobile Industry is Changing

The revolution in communications that has put a phone in four billion pockets[1] started in December 1982 when a group of European engineers and regulators called the “Groupe Spécial Mobile” (GSM) met in Stockholm to specify a pan-European mobile phone system. In the subsequent 27 years, not only has their original intention been realised, probably beyond their wildest dreams, but the technology itself has evolved through three radically different generations driven by the forces of silicon technology, the move to IP networks, and the human need to connect. The GSM system was intended to replace the national “first generation” standards that had originated in the USA and Scandinavia with a system that would allow European citizens to use their phones anywhere in the single market.

As a result of those efforts, an enormous new industry has grown up and flourished, but is now facing saturation in its markets, flat or even declining revenues, and consolidation at every point in the supply chain. In less than three decades, mobile communications has become an almost universal utility.

The original GSM standard evolved to provide very efficient voice communications and messaging, and was later enhanced to carry always-on packet data. Despite its seeming complexity it permitted the manufacture of extremely cheap and small handsets, and these drove huge market growth, initially in Europe, then in the USA and the rest of the world. Even now in emerging economies GSM phones are helping millions of people struggling to live better. The growth of multi-media communications in the 1990s; and the possibilities created by auctioning newly-vacated spectrum; led to the introduction of third-generation (3G) systems such as UMTS and CDMA2000 in the early years of the new century, which were supposed to make the Internet mobile. These systems have not fully met that promise (though still being rolled out in many countries and serving some 800 million users), and now the industry is in the early stages of moving to LTE and LTE Advanced systems which should finally deliver nearly universal, wireless, broadband IP voice and data connections to anyone, anywhere, on a small pocket terminal.

It’s instructive to compare this history with the PC industry – whilst the latter has benefitted from, and in many ways driven, the same silicon technology evolution, the basics of how a computer operates have not changed significantly during a period where mobile communications has moved from largely analogue circuit-oriented FDMA through digital TDMA to CDMA and now the OFDMA and SC-FDMA packet-oriented air interfaces which will be used in LTE. Even when the new air-interfaces are introduced the mobile operators have demanded compatibility with networks already installed, so every 3G phone today operates also on the legacy GSM networks; and LTE devices will have to operate on 3G and GSM. It will be many years, if ever, before the legacy networks and their frequencies are converted to the new standards and single-mode terminals become the norm.

In the early years of GSM, a number of very large established companies which had the scale to invest dominated the development of not just the radio technology itself but the necessary VLSI devices. Once the market took off many new entrants appeared making mobile phone chips,

---

[1] Source: GSMA market data as at end 2nd quarter 2009

attracted by the huge scale (a factor of at least ten times larger than the PC market), and this in turn encouraged many new handset makers especially in Asia. Today though the situation is changing and simplifying through three interconnected factors.

- Many mobile markets have reached saturation, where every potential customer has at least one and in some cases two or three phones. The service providers effectively can't add any new subscriptions but have to fight for market share; and the intense competition between them is forcing down tariffs. Customers have proven remarkably unwilling to pay much extra for mobile broadband even though its usage is increasing rapidly, to the extent that mobile networks are running out of capacity. Capacity can only be increased by investing in more infrastructure. Faced with flat or even declining revenues, and needing to spend more on their networks and new spectrum, mobile carriers are starting to coalesce so they can share infrastructure and investment.
- As the standards have developed the investment needed in terminal software has increased hugely, especially as it must support multi-mode and multi-media communications and applications. The scale of the industry also favours large handset suppliers who can support major operators and their customers. In many countries the operators have heavily subsidised handset sales which has made their business plans very sensitive to the device cost, and this pushes down the handset makers' margins. These pressures favour the largest and most efficient manufacturers and are leading to consolidation in the handset supply chain.
- New standards affect the radio modem most of all, and suppliers of silicon devices have needed to invest, not just in the right silicon processes but also in designing the modem itself. The increasing complexity of the standards is making it very difficult for a merchant silicon supplier to maintain the necessary investment in the modem technology, especially as the timescales for introducing a major new variant of a standard are significantly longer than the normal payback cycle for investing in a major new chip design. Many of the new entrants to the cellular market from the early 2000's have now exited and left three main suppliers of 3G – capable baseband devices, Infineon, ST – Ericsson, and Qualcomm; and a small number of GSM-only suppliers.

Against this background the industry now faces another major change in technology as LTE and LTE-A; and beyond them ITU-Advanced (which may at last be a true global standard) are introduced. As markets saturate; and the capability of the wireless technology will equal or exceed that of the fixed access networks that most people can connect to; we can predict that in the next decade wireless communications will move to become an essential utility; with utility – type returns on investment in products and infrastructure. This will affect how every part of the industry does business. In this paper we consider how it affects the design of new terminals in particular and why a new generation of signal processor devices is needed to streamline modem development and protect investment in the new standards.

Standards Evolution

Though other technologies have their place, there can be no doubt that the growth of the global market has been driven by GSM and the related standards developed by, first the GSM; then ETSI and the Third Generation Partnership Project (3GPP) that subsumed the GSM standards in 1999. It took from 1982 to 1991 for GSM to be specified and the first system launched. UMTS specification started in the late 90s; 3GPP was founded in 1998 and its first standards published a year later. Figure 1 shows the general roadmap of the 3GPP specifications.

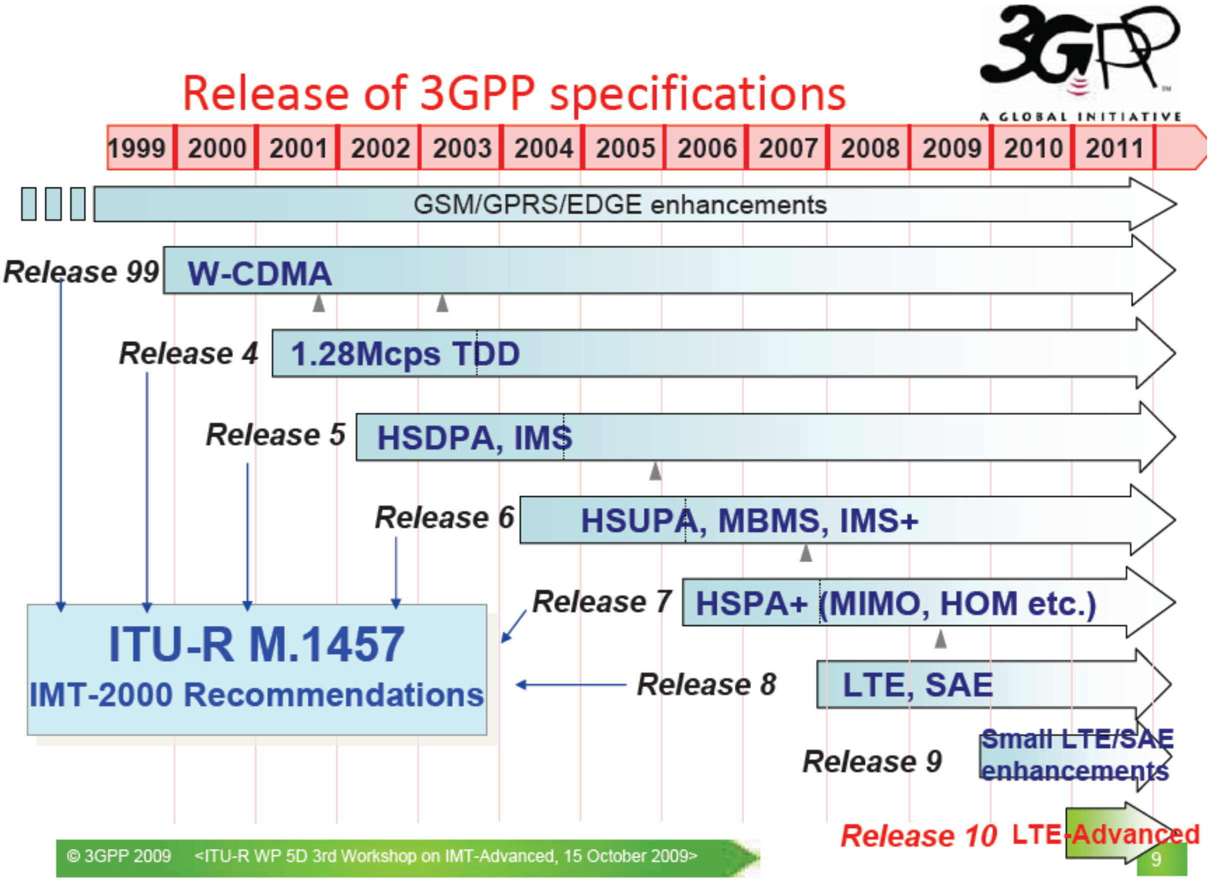

Figure 1: 3GPP standards roadmap[2]

The GSM family of standards continues to evolve slowly. GSM provides an economical basic voice and data service using low frequencies with good wide-area coverage. A GSM phone will connect almost anywhere its user travels over the globe. GSM networks will continue to operate even as LTE is deployed, and the system needs to evolve to inter-operate with the new standards. One of the applications that may become very important on GSM is “machine-to-machine” communications (M2M). M2M does not require very high bandwidth but does require low equipment costs and excellent coverage.

From Release 5 to Release 7, the UMTS standard was enhanced with key features under the “HSPA” (and HSPA+) umbrella that add optimised packet data transfer based on high-speed shared channels,

---

[2] Source: 3GPP

with higher-order radio modulation and MIMO to optimise data rate in HSPA+. These features are being progressively deployed today, and will provide nearly equivalent performance to LTE at least for high-grade terminals used close to the base station.

Published in early 2008, Release 8 defined LTE; and a new all-IP system architecture called SAE. LTE uses a radically different air interface designed to carry IP efficiently and optimise system performance further for users at the cell edge. LTE systems will be deployed from late 2010. LTE will grow in importance and become dominant, and will be deployed in new frequency bands such as at 2600MHz and 800 MHz[3] as well as those in which GSM and HSPA currently operate.

The move to LTE is driven not just by broadband services. Essentially, the very fat IP “pipes” that LTE provides will optimise statistical multiplexing efficiency for the whole range of services from simple voice and SMS through multi-media streaming to broadband interactive applications, and allow the operators to make the most efficient use of their spectrum.

Another factor that has to be considered is that as broadband mobile networks are rolled out, LTE is not the only standard of interest. WiMax (IEEE 802.16m) is a mobile broadband standard also targeted at the 2.6 GHz frequency band, with many similarities to the LTE air interface. WiMax networks are being deployed in many territories and will compete with LTE at least for high-speed data applications. In Korea a similar system called WiBro is also operating. Certain classes of mobile device such as laptop computers may need wireless modems that can operate on many unrelated standards, such as HSPA+/LTE, WiMax and WiBro.

It may seem that standards will continue to evolve rapidly judging by the changes seen in the last decade, from GSM through WCDMA and HSPA to LTE. An alternative view however is that, as the importance of IP as a universal information transport protocol for the future has become evident; and as more has been learned about how to manage high-bandwidth data transmission over the air; WCDMA has been recognised as having shortcomings and LTE, and LTE Advanced, will be around for the foreseeable future and the technology will progressively be deployed in those frequency bands where 2G and 3G operate today.

---

[3] The so-called “digital dividend” band in Europe.

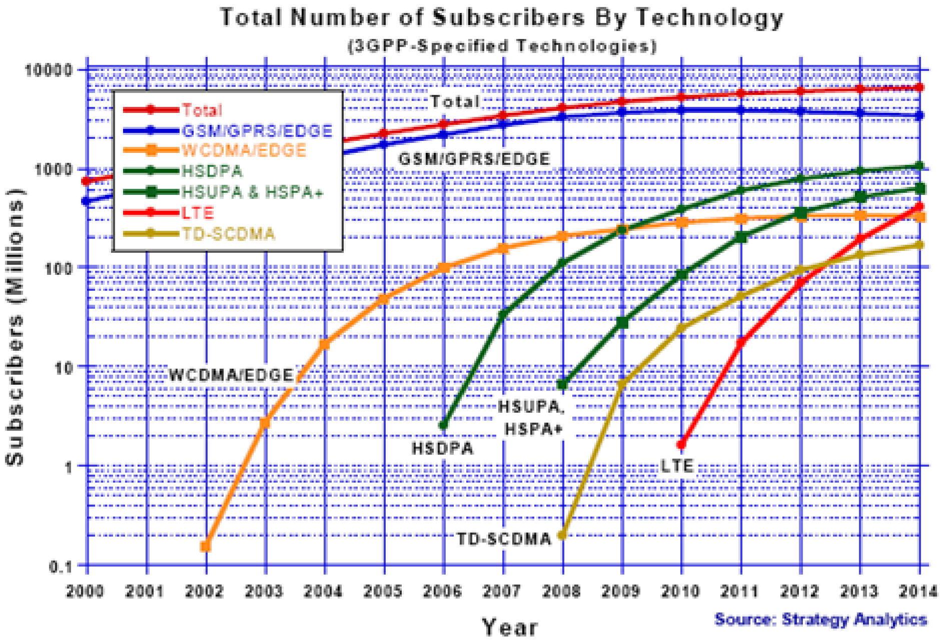

Figure 2: Projected market growth

Figure 2 shows how the global market is expected to evolve into the next decade. Noting that the vertical scale is logarithmic, the exponential growth period for total subscriber numbers of the 1990's and early 2000's is over, and GSM and its derivatives are expected to dominate for some years to come. LTE is expected to start rolling out in 2010 and then grow rapidly and will become the dominant standard. One clear message is that the market for personal mobile subscriptions has become saturated[4].

Terminal modem development

Modern cellular handsets are highly-optimised communicating computer systems which combine many functions with small size and very low power consumption. It is only possible to make them by using the most advanced VLSI technology, and highly-optimised application-specific integrated circuits (ASICs) have to be designed to provide the functionality needed in a small space with low power consumption. A more detailed discussion of modem design approaches has been given in a related White Paper from Cognovo[5], so the information below is a summary.

First-generation GSM devices used a combination of control & signal processors and custom hardware (for e.g. Viterbi & crypto processing). As processors became more powerful more functions were absorbed into software and today it is possible to realise a complete EGPRS modem, including protocol software, in code running on a DSP core.

---

[4] Note however that if M2M applications take off the potential growth could be very large. As noted above though, M2M may not require large bandwidth, and its applications are very sensitive both to device costs and tariffs.

[5] http://www.cognovo.com/Documents/Managing%20the%20multimode%20architecture%20evolution.pdf

With Wideband CDMA the processing requirements were beyond the capability of DSP devices of reasonable power consumption for handsets, so much of the signal processing was implemented in around 1.5 million gates of dedicated hardware and associated memory. The rapid evolution of WCDMA to HSPA and HSPA+, which are considerably more complex, has meant that this approach persists and an HSPA modem contains a powerful digital signal processor core; one or more microprocessor cores for communications protocols and applications; memory systems to support these processors; and up to 2.5 million gates of dedicated signal-processing logic with associated data memory

The move to LTE adds the requirement for yet more dedicated hardware, more memory, and even higher performance processors, and it is estimated that 3 – 4 million extra gates are needed for LTE. Figure 3 illustrates this increase in estimated processing capability with system generation. Essentially since the introduction of UMTS the on-air gross bitrates and the required processing capability have increased exponentially with generation. LTE approximately doubles the requirement compared to HSPA+, and LTE-Advanced will need 5 – 10 times more than LTE.

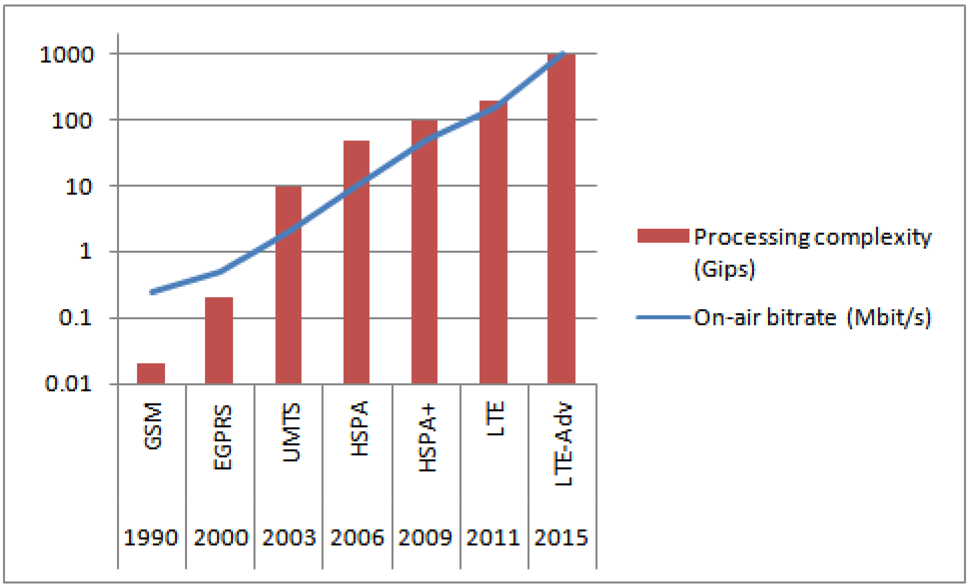

Figure 3: Processing complexity of different cellular systems

Modem design flow

Figure 4 shows a typical design flow for a modem realised as a combination of processor cores and dedicated hardware.

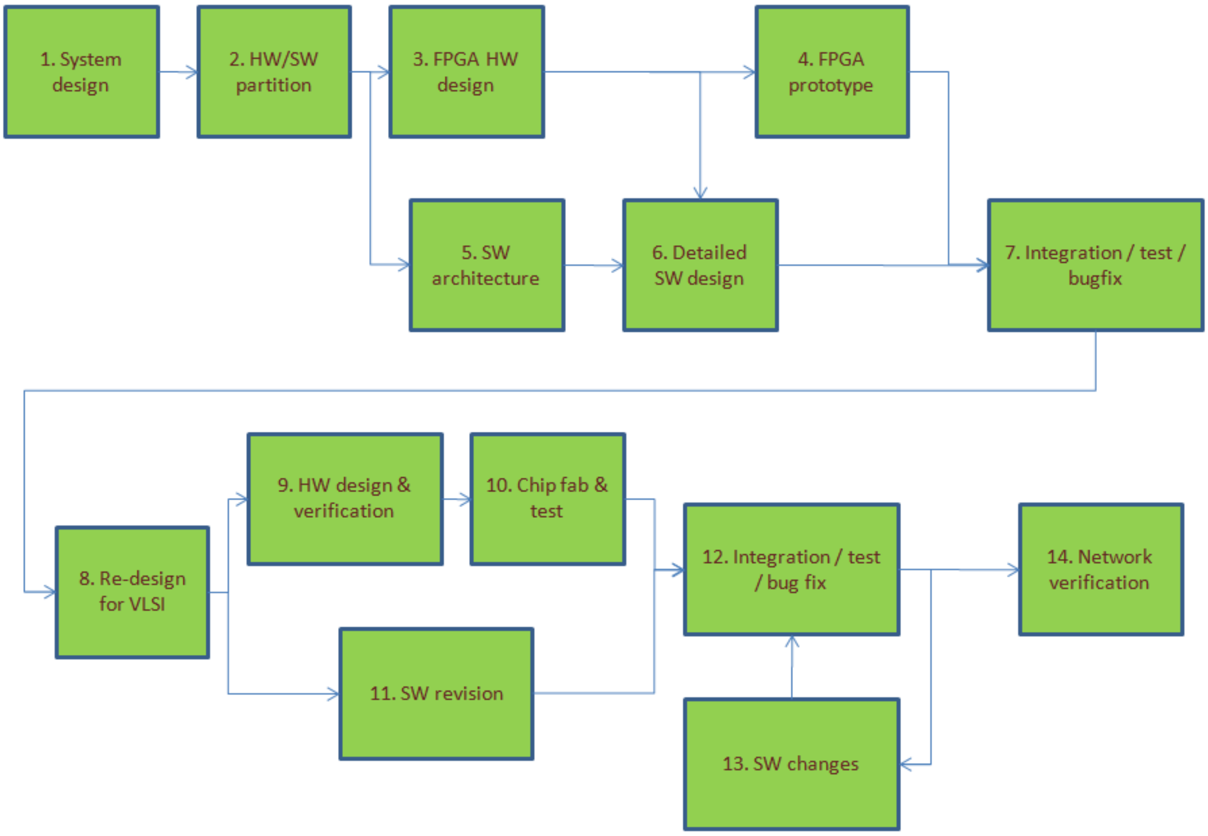

Figure 4: Conventional design flow

At the start the overall system design has to be determined and whether particular parts should be realised in dedicated hardware or software. This determination is critical, since complex functions such as algorithms realised in hardware get frozen at quite an early stage and can't easily be changed later if there is a problem; whilst too much processing by software increases the capability required in the DSP core and may overload the core selected. Once the overall system is designed, an initial prototype is required to prove the overall system functionality and this is usually realised as a combination of custom hardware realised in FPGAs and software on the selected processor cores. However the design of both is heavily conditioned by the hardware platform since FPGAs come in predetermined sizes and have speed and interface limitations. The detailed software design can only be started once the FPGA has been architected and aspects such as interfaces and memories decided on.

Even with an FPGA platform there are aspects of system operation that cannot be proved in detail. It is rarely possible to run the FPGA fast enough to allow real-time operation of the system. For LTE this will be a real problem since either the whole system has to be run at "scale speed" in which case it cannot easily be tested against a real LTE test system or network; or the tests will be limited in functionality (for example in RF bandwidth and/or application data rate). The first of these is the biggest problem since network interoperability is the final arbiter of design success. There are also aspects of system operation such as Idle Mode operation or various data use cases that cannot be fully emulated in an FPGA. Advances in FPGAs improve the situation, but these tend to focus on increasing size rather than speed; and in some ways this introduces more problems since the designer has to devote time to solving issues like clock distribution on a large FPGA which does not contribute to the final ASIC implementation. Thus whilst the integration and test of the FPGA system is important, it is not the end of the story.

Once the FPGA system is proven, design of the final chip can start with more confidence. However, a significant amount of re-design of the hardware is required for an ASIC implementation. The component libraries will need to be adapted; interface and bus design will be different; power management features (for example switching off parts of the chip not in use at any given time) have to be designed; and the approach to on-chip memory will be very different. All of these changes can only be verified by simulation at this stage. At the same time the software will have to be updated to run on the final ASIC because of issues such as interfacing, memory management, and power control.

Once the chip design is complete a test device can be fabricated. Only once this is available and tested (including the additional features that can't be tested in the FPGA) can proper integration and test of the software begin. It is now possible to test at full speed/bandwidth, and to test aspects such as Idle Mode operation, that cannot be done on the FPGA platform.

It is difficult to provide good visibility of signal flow in dedicated hardware, so problem diagnosis can be hard. Once problems are found, they cannot be fixed in the hardware without a chip re-spin; so often somewhat imperfect workarounds have to be adopted in the software. Once the ASIC and software are working as designed, it is possible to start testing against network infrastructure which is the actual goal of development. Unexpected problems are frequently found in certain network configurations – for example because of the way the infrastructure implements optional aspects of the standard; or through radio interference issues. At this point the only option usually is to fix problems through software workarounds since critical algorithms are implemented in hardware, and the protocol software undergoes a process of "hardening" as handsets are developed and tested on real networks.

This whole process is quite clumsy and likely to be error-prone. The fact that it involves two hardware design steps makes it longer; it is difficult and can take a long time to find and fix hardware errors; and the necessity to re-spin the chip if there is an error which can't be worked around in software will insert additional delay and considerable extra cost.

Another issue to be considered is the evolution to LTE Advanced and then IMT Advanced. Even though these are some years away, LTE-A is designed as a scaling up of LTE; and with the large cost and complexity of the standard development model it will be vital to preserve as much as possible of the development investment from LTE.

We can see that the move to LTE introduces a need to more than double the complexity of the dedicated hardware in a modem, and development of LTE devices will be very difficult using the approach above. With the move to new silicon process nodes, the cost of the initial development of a chip; and its re-spin if there are errors; is also becoming very high.

Semiconductor process development

Wireless handsets are only possible because of advances in VLSI technology, and today's HSPA and HSPA+ handset baseband chips use 90nm or 65nm silicon geometries. Each new silicon process node greatly increases the possible complexity of a reasonably-priced device, its processing speed, and reduces power consumption. LTE and LTE-A will require 32nm devices to be viable.

As each process node is introduced, the cost of designing a device and of fabricating a wafer also increases.

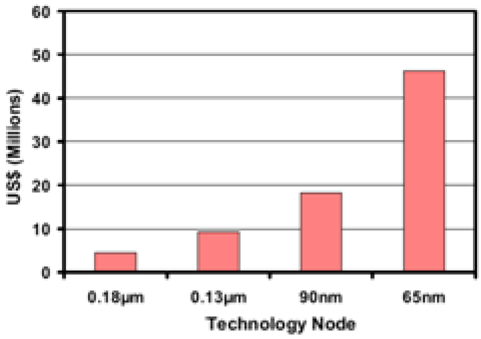

Figure 5: Design costs for a new VLSI device[6]

Figure 5 shows that the cost of designing an average complexity device using "bleeding edge" VLSI technology increases rapidly with process node. Much of this cost increase is incurred in the much more complex design verification which is needed as the device density increases; and the foundrys' amortisation of initial set-up costs. Later on in the life-cycle of a given process node, the cost declines, but the latest cellular standards are only made viable by the new process so they inevitably get designed early on when a process is just made available. The initial costs of wafer fabrication also rise sharply as the feature size decreases.

The combination of these factors means that the up-front investment to develop a new baseband devices are increasing exponentially Of course, the cost per device reduces significantly, but it becomes increasingly necessary to achieve large sales to be profitable overall. The business risk is increasing.

The modem design singularity

We have seen that the complexity of wireless modems is growing exponentially with the introduction of LTE and LTE-A; and also that the costs of designing a new device in the latest process technology is rising exponentially. At the same time, growth in the cellular market has slowed down and it is becoming ever more important to minimise development costs and amortise them across as many products as possible. This will be important for devices for handsets, and also for the emerging femtocell market sector which is growing in importance for 3G and many expect to be a core component in LTE deployments.

---

[6] Source: Chang (TSMC); *IEEE International Solid-State Circuits Conference*; 2007.

The large costs of designing the next generation baseband device and of having to re-spin the hardware if needed; plus the relatively long time span for LTE to dominate the market; mean that the way that cellular modems are designed and made must change. Ideally a modem should be designed using a fully programmable, multi-mode hardware platform – the **Software Defined Modem**. For this paradigm shift to be possible a number of factors have to come together.

- Most importantly, the processor has to have enough raw processing power to undertake the complete modem processing task (though it may be appropriate to provide specific accelerators for generic functions such as turbo decoding and FFT).
- For handsets, where battery life and cost are critical, raw MIPs have to be combined with power and silicon area efficiency.
- A development platform with the full performance of the processor is needed, so that full speed testing of all the functions can be done as early in the development cycle as possible.
- The processor system has to be specified to meet the special needs of cellular modem development. This encompasses aspects such as power management; fine-grained management and scheduling of processor tasks (especially for multi-mode operation); and a software framework enabling clean separation and layering of protocols.
- The development of the processing platform and its evolution has to be decoupled from the modem system development and evolution, so that the chip development can be done largely by silicon companies without detailed modem design expertise, and the chip can benefit from improvements in process technology.
- A consistent platform architecture roadmap is needed across generations to enable design re-use as LTE evolves to LTE-A, and for addressing other modes that may be needed such as WiMax and legacy air interfaces such as HSPA+.
- The tool-chain has to allow designs to move smoothly from system design right through to terminal development and testing. For the critical signal-processing functions, they must allow an easy and automated transition from industry-standard algorithm development tools such as Matlab or Simulink to optimised running code. Development tools and test tools should use a consistent schema and provide visibility of events and signals at all layers; and at all stages of product development from initial R&D through design verification to field trouble-shooting.
- Platform software and hardware architecture must allow import of legacy software modules; and if appropriate hardware modules; to add legacy modes of operation at reasonable cost and preserving the operational “hardening” built up through years of exposure to real networks.

Figure 6 shows the streamlined design flow for a Software Defined Modem.

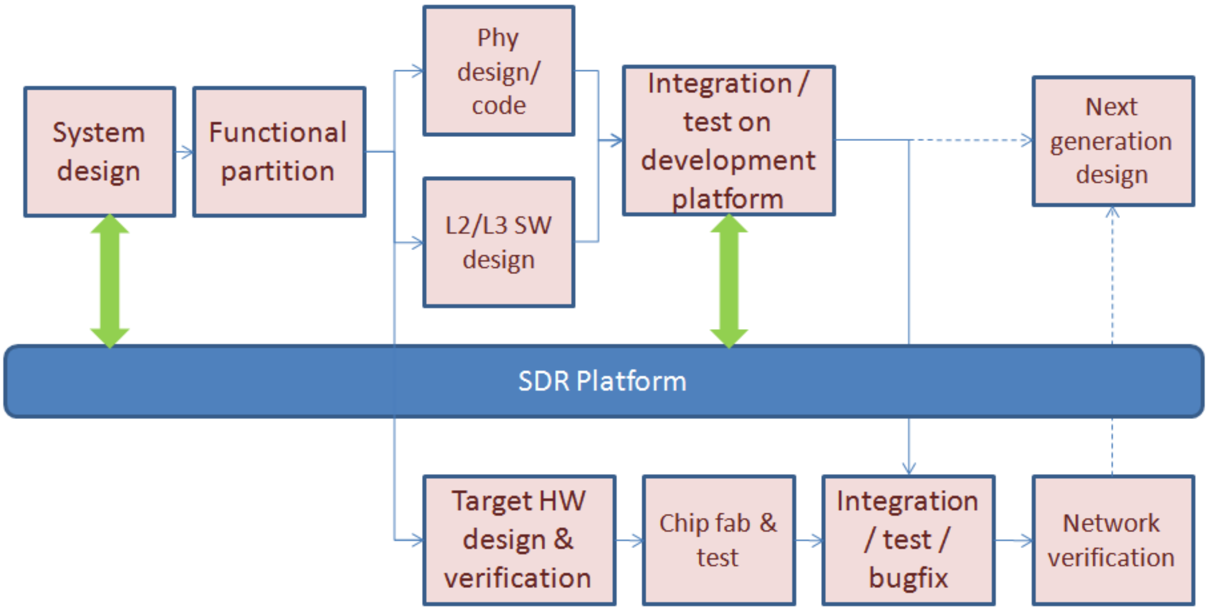

Figure 6: SDM design flow

With the SDM, a development platform is available which is largely identical to the final device except for aspects such as external interfaces and exact memory sizes, and which is capable of operating at the full speed required. Now, all of the design effort can be targeted at the final device rather than having the intermediate FPGA platform; and prototyping and testing can be done on a platform with the full capability of the final device. Moreover, because the modem is defined in software, changes can be made to fix significant design errors without hardware changes, even once the device has entered service. Development of the actual production device can be started much earlier in the design cycle, in parallel with the detailed modem design; and it is much less likely that hardware re-spin will be needed. Testing can be done on real network test equipment or actual networks, for all the use cases of interest. The detailed behaviour of all layers of the system can be simulated or observed at each stage of development, allowing much more direct comparison of field experience with the initial design for example.

The Cognovo Vector Signal Processor (VSP) is a licensable core suitable for realisation in 45 and 32 nm silicon geometries with enough raw power combined with low power and small on-die size to meet the LTE processing challenge. As well as the core design, Cognovo supply a software framework and design tool chain to meet the other requirements. The VSP scalable architecture and design roadmap will maintain a consistent software environment and operational behaviour, and allow devices for LTE-A and IMT Advance to inherit much of the design done for LTE to minimise the investment needed as these standards reach service.

Conclusion

When mobile phones started their phenomenal market growth in Europe, Japan and the USA three decades ago, they entered a market where virtually every household had a fixed telephone. In those early days some people wondered if there could be a consumer market for a mobile phone. But today we are rapidly reaching a state where virtually every individual in the world has a mobile phone, and there are complete generations who have never known a world without them. In the next decade the network technology will stabilise around a system providing ubiquitous broadband internet access. With markets saturating and tariffs forced down through competition, mobile communications is becoming a utility just like the fixed phone. Silicon and processor technology have caught up with the standards, and the Software Defined Modem will be the long-term foundation for the handset industry.