Digitization and Edge Processing Redefine the Roles of Semiconductor and Packaging Technologies for Future Defense Platforms

John Cowles

Analog Devices, 14320 SW Jenkins Rd.
Beaverton, OR 97005 USA
john.cowles@analog.com

Keywords: Phased Arrays, Electronic Warfare, Digitization, Smart Partition, Heterogeneous Integration

Abstract

Requirements for reduced size, weight and power (SWAP) and increased autonomy and multi-mission flexibility of future defense platforms have brought the digitization and decision making closer to the antennas. Integrating diverse semiconductor technologies and passive components into a single package are fueling changes in how signal chains are architected. Smart partitioning to achieve the system requirements with a reconfigurable platform mindset is now more critical than individually optimizing stand-alone components.

INTRODUCTION

These are unprecedented times where electronics impacts all aspects of daily life. The confluence of computational power, memory storage, networking and cloud services, radio technology and digitization are revolutionizing transportation, healthcare, factories, communications and entertainment. Resulting efficiencies not only reduce environmental impacts but also drive down the cost of ownership. The electronic content of just about every platform is increasing dramatically in proportion to the structural content. For example, while the projected broad market growth of semiconductors from 2016 to 2026 is estimated at 7.5% CAGR, automotive semiconductors are expected to grow at 11.5% pace in the same time frame. Electronic systems managing the motor drive, battery health, infotainment systems and wider adoption of electric vehicles are driving this spend. It should be noted that electric vehicles consume twice the value in semiconductors as internal combustion engines.

Aerospace and Defense is experiencing a similar trend. It is estimated that worldwide defense budgets are approaching $1.3T ($773B in US) in 2023 due to geopolitical tensions and desire to home-grow capabilities. A greater fraction of the budget is allocated to electronics, cyber security and artificial intelligence, not to mention adjacent spend to on-shore semiconductor manufacturing critical to national security. While military budgets are growing at roughly 4% CAGR, spend on electronics is estimated to be growing at twice this rate. Ground, sea, air and space-based platforms are carrying increasingly more intelligent sensor nodes for spectral dominance, navigation and threat detection. Complementing defense applications is the deployment of massive commercial low earth orbit (LEO) constellations. These smaller satellites pack more electronics in to a smaller volume compared to the larger geosynchronous earth orbit (GEO) satellites of the past.

Fig. 1. Today's environment in the battlefield is increasingly complex and dominated by electromagnetic waves that carry communication and radar signals. The electronic content of each platform is increasing to support dominance of this spectrum.

The increased sophistication of military threats and their faster deployments are putting pressure on existing platforms and capabilities to respond. Rather than developing brand new combat platforms over long cycles, the DoD is upgrading existing ones with newer radar, electronic warfare (EW) and communications systems, sometimes as consolidated multi-function systems. This proliferation of on-board sensors, the deployment of electronically steered phased arrays (AESA) and the widening of the contested spectrum together generate massive amounts of data in need of processing. Cloud computing is not only prohibitive in terms of high bandwidth connectivity in hostile environments but also introduces unacceptable latency to fast decisions and cybersecurity risks. The obvious solution is to embed the digitization, compute and decision making locally into the sensors with RF, creating self-contained Antenna to Insights system. FPGAs, AI hardware accelerators, data converters, digital signal processing must be integrated with RF components and antennas, all supported by efficient power management. Considering the system holistically offers the possibility of repartitioning the signal chain by technology to achieve the challenging size, weight and power (SWAP) requirements needed to deploy these sensors effectively.

EARLY EXAMPLES OF REPARTITION

The first examples where the repartition between semiconductor technologies occurred was around RF sections of early mobile phone and base-stations (BTS). In the early days, the entire radio was implemented in GaAs MESFET and then pHEMT technologies, often as discrete building blocks like mixers, attenuators, power amps (PAs) and low noise amps (LNAs). Frequency sources were often silicon-based synthesizers with specialized GaAs varactors for tuning the voltage-controlled oscillator (VCO). In time SiGe and then RF CMOS reached adequate RF performance levels to replace most of the radio, except for the PAs and, in some cases, LNAs.

The level of integration and reconfigurability of silicon-based solutions also enabled digitally assisted RF where imperfections in the RF stage could be sensed and corrected in the digital domain. Techniques like quadrature error correction, noise cancellation and digital predistortion (DPD) became standard and led to significant improvements in channel density with even better performance. Thanks to digital techniques, mobile links today with software-defined radios (SDRs) cover wider, disconnected frequency bands, cope with self-interference and take advantage of multi-path effects using massive MIMO architectures.

Silicon radios did not spell the end of GaAs technologies, which are needed at the very front-end to generate power and provide receiver sensitivity. In fact, GaAs and newer GaN technologies have evolved to work symbiotically with the silicon radios to achieve higher overall efficiencies for extended battery life in the mobile units and lower the cost of ownership (power consumption) in base stations. A key aspect of PAs today is how linear and efficient they remain over output-power back-off (OBO) from peak and with complex modulation schemes. Several clever architectural techniques have been resurrected such as Envelope Tracking and Doherty topologies to address the efficiency vs OBO dilemma. How well these amplifying systems respond to DPD algorithms is more critical than how linear and efficient they are without any correction, (“naked”).

A related realignment between silicon and GaAs technologies happened in the microwave backhaul for cellular networks. These microwave point-to-point systems connect base stations that do not have fiber connections to the wired network. Along with satellite broadcast systems, these were the first true commercial applications of microwaves for telecommunications. Microwave backhaul covers frequency bands ranging from 6 GHz to 85 GHz. It typically consists of a microwave outdoor unit (ODU) near the antenna that converts the microwave channel to an intermediate frequency (IF) in the 100’s of MHz. The indoor unit (IDU) on the ground digitizes the IF for delivery to the wired network or to a second microwave link hop. While the IDU has naturally been in silicon, the ODU has evolved from frequency specific discrete GaAs signal chains to more highly integrated wideband silicon SDRs complemented by GaAs and GaN front-ends.

Improved RF performance in silicon combined with digital signal processing overcame limitations in operating frequency, signal-to-noise ratio (SNR), phase noise and wideband quadrature accuracy. Today these microwave links can support >2Gbps data rates using >1024 QAM constellations, with dual polarization over 100MHz channels. Again, GaAs and GaN front-ends are indispensable to achieve the needed link margins over several miles but had to adapt to the silicon radios to achieve the needed transmit output power and overall receiver NF. Together, the SDRs in silicon and redesigned GaAs/GaN front-ends achieved higher capacity at a lower SWAP than before.

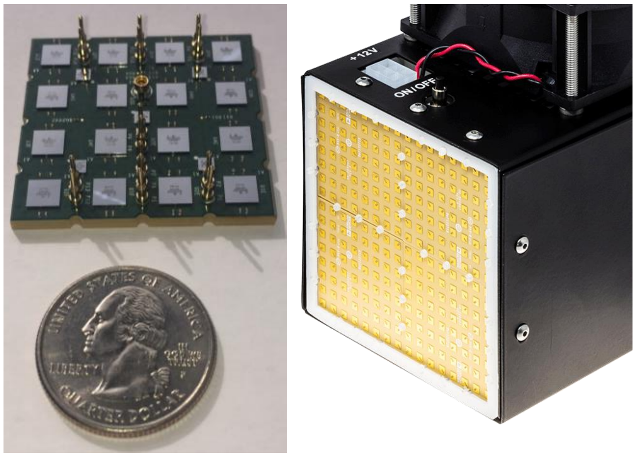

Fig. 2. AESAs have proliferated in 5G, LEO/GEO satellite communications and radars thanks to RF silicon integration, efficient digital processing and efficient power amplifiers. Planar arrays are now common up to 40 GHz while further advances in packaging will enable arrays beyond 100 GHz.

PHASED ARRAYS MULTIPLY IT UP

Phased arrays have existed since the first half of the 20th century as a method to direct and receive energy preferentially. Thanks to improved semiconductor RF performance and integration, electromagnetic (EM) modeling and digital processing, they are becoming practical and scalable. Multiplexing in time, frequency and code have reached the limits set by Shannon’s Capacity Theorem in terms of information transmission. The ability to spatially multiplex dynamically opens another dimension to exploit for more capacity per area. While phased array radars for military applications have been in operation for decades, commercial applications are just starting to deploy with 5G at 28 and 38

GHz and with LEO satcom networks at Ku and Ka-bands. There is an interesting merging of massive MIMO concepts and phased array technology, especially when element level digital beamforming is considered.

The nature of phased array is to create an aperture of regularly spaced transmitters and receivers, usually spaced no more than half of the operating wavelength to avoid what are called grating lobes (the spatial analog of aliasing in time sampling). By individually controlling the amplitude and phase characteristics of each element, energy can be directed to and received from prescribed directions. Arrays of 256 or even 2000 elements that replace a single large transmitter with mechanical steering demonstrate one reason why electronic content keeps increasing. The larger arrays create sharper beams and the possibility of "fitting" more beams simultaneously into a hemisphere.

More directive beams improve the sensitivity and resolution of communications systems and radars. With the use of clever analog or digital signal processing, multiple independent beams can be created simultaneously. A 16-beam system can point 16 different channels in different directions for a satellite or cellular network or scan the horizon 16x faster for a radar. This type of flexibility is just as important as being able to focus and steer energy from one location to another. To create multiple beam systems, the number of amplitude and phase control blocks must increase proportionately. This increase has significant implications to partitioning by semiconductor technology and by physical implementation.

The classic embodiment of phased arrays necessarily relied on GaAs signal chains that included phase shifter, variable attenuators, driver amps, PAs, LNAs, switches and sometimes limiters. Each antenna element was backed by one of these transmit-receive modules (TRMs) either as discrete implementations or as integrated monolithic microwave integrated circuits (MMIC). All the received signals from the TRMs were combined before a single down-converting mixers to an IF for further processing. Analogously, after the up-conversion mixer, the signal was split to all the TRMs. This is an example of RF beamforming where all the antenna signals have their own gain and phase weights applied between the mixers and the antennas.

The quality of the beam shapes and spatial rejection depend strongly on the accuracy of the amplitude and phase steps and their matching from channel to channel. These early implementations required extensive calibration because channel to channel variation was large and, within a channel, gain changes induced unwanted phase changes and vice versa. GaAs-based gain and phase TRMs struggled to fit the half-wavelength spacing at higher frequencies and had to be implemented as blades that extended behind each antenna element.

Once SiGe and Silicon-on-Insulator (SOI) technologies matured in the mid 2000's, so-called Beam-Former Integrated Circuits (BFICs) were developed to support X-band radar, millimeter-wave 5G and Ka/Ku band satcom. These multi-channel ICs deliver excellent matching and absolute accuracy with minimal contamination between gain and phase changes. They fit behind a set of antenna elements in a planar format rather than blade orientation. Furthermore, they can store gain/phase states in memory and provide controls to the front-end components, when still needed. The density possible with silicon even allows the doubling and even quadrupling of gain and phase control paths per element to support dual polarization and/or multiple beams.

Today there is another seismic shift in phased array architectures thanks to direct sampling and synthesis by data converters up to X-band and even Ku/Ka-bands. Digitization at each element transforms the gain and phase changes from the RF domain to the digital domain. Digital beamforming not only allows more flexibility in setting gain and phase weights to each element path but also supports an arbitrary number of beams and can generate true time delays as opposed to phase shifts to avoid beam squint. This transformation is happening today as sampling rates increase, the power dissipation in the digital portion drops and digitization comes closer to the antennas.

With silicon-based RF and digital beamforming, TRMs become greatly simplified to a PA, LNA and a switch or circulator depending on whether the system is time or frequency division duplexed. While the digitization portion of the system becomes a reusable platform, the TRMs give the system its personality. Exactly as with the microwave links discussed earlier, the GaAs or GaN PA plus driver must raise the output level to the required power at the antenna. The LNA needs its NF and gain to offset the NF of the rest of the receiver. Because of the large number of channels and the tight spacing imposed by the half wavelength spacing, power efficiency, size and ability to co-package become critical aspects to enable a planar, even conformal form-factor.

ELECTRONIC WARFARE DEVOURS THE SPECTRUM

While phased arrays process a large number of channels operating over modest RF bandwidths from L-band to Ka-band, Electronic Warfare (EW) systems are looking to process wide swaths of bandwidth at a time over a 2-50 GHz range. EW systems come in different flavors, sometimes just scanning the spectrum, other times generating signals that confuse, jam or spoof an enemy system. This contest for spectral dominance is at the forefront of modern warfare where the key goals are to scan, identify and react to the environment as quickly as possible. This translates into digitizing wider bandwidths at a shot, spot signatures of interest (increasingly via machine learning) and take appropriate action which could be synthesizing and transmitting signals to counter the threat. At the same time, EW systems need to be miniaturized so they can be carried by soldiers, mounted on aircraft wings and so on along with other sensors.

The key technology enabler is the ability to digitize wider bandwidths over a wider spectrum. A typical deployed system today is limited to 100's of MHz instantaneous bandwidth. To cover just the 2-18 GHz range, a number of mixer stages and narrow band filter banks are needed to channelize the spectrum. There is a major trade-off between system agility and SWAP since faster, parallel processing requires multiple signal paths with bulky filters, numerous amplifiers to balance loss, local oscillators for mixers, and clocks for different data converters.

The same direct sampling and synthesis technology that was mentioned in Digital Beamforming for Phased Arrays is also repartitioning EW platforms. For EW, the ability to digitize as much as 4 GHz at a time while maintaining >50 dB of spurious free dynamic range greatly reduces the system complexity. Furthermore, the incorporation of critical DSP like digital up- and down-conversion, numerically controlled oscillators, digital filters, decimation and interpolation, among others provides flexibility in the sub-channelization and signature identification. On the transmit side, the DSP can produce arbitrary waveforms directly at frequency in the same way that Direct Digital Synthesis (DDS) operates.

The impact of wide-band digitization closer to the antennas is that the RF signal chains need to support wider bandwidths. These architectures require wider band amplifiers, switches and mixers. The usual narrow-band tunable filters can now be fixed or at most course-tunable. Wide bandwidth signal paths introduce the importance of second order distortion, possibly favoring differential implementations. The entire level planning needs to be reconsidered since the RF paths are greatly simplified. More importantly, the RF signal chains need to miniaturize either by die-level integration on laminates or by monolithic integration when appropriate.

BEYOND MOORE

For several decades, Moore's Law of CMOS scaling continued to fuel System-on-Chip (SoC) integration where all functions RF, analog, digital, power management would be implemented monolithically. SDRs took advantage of digital techniques not only to process data but to correct analog imperfections in the signal chain, all inside the chip. However, the benefits of scaling analog/mixed-signal and RF into the most advanced CMOS nodes are questionable, creating development time and cost challenges for the SoC's paradigm. Heterogeneous integration is an alternative that enables the industry to go Beyond Moore by rationally disaggregating functions into reusable blocks, possibly in different technologies, while maintaining the interconnect integrity between them.

The evolution of multi-technology packaging goes back to the 1980's with multi-chip modules (MCMs), followed by systems-in-package (SIPs). Today the focus is on chiplets and 3-dimensional heterogeneous integration (3DHI) that allow different die to be combined compactly. Advanced packaging must support high speed data links between data-converters, digital processors like FPGAs and high bandwidth memory as well as RF paths between antennas, front-ends and radios.

Packing more functionality into smaller form factors creates new challenges. Thermal dissipation becomes more difficult since natural heat paths are obstructed. Stacking layers with different thermal, electrical and mechanical properties requires detailed multi-physics simulation capabilities to model and predict heat flow, stresses and signal integrity between blocks. Going Beyond Moore requires an eco-system of advanced modeling and packaging to complement smart systems partitioning and expertise in different semiconductor technologies.

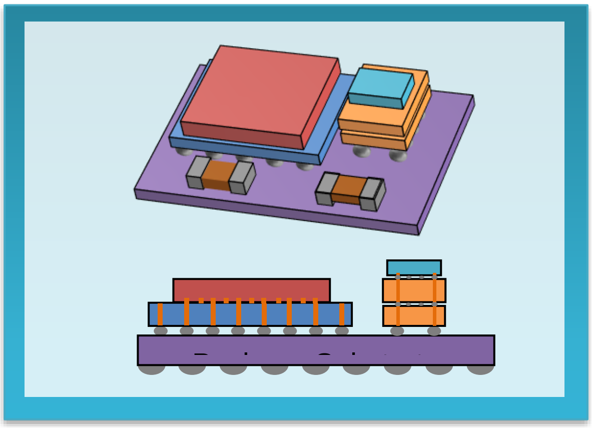

Fig. 3. Heterogeneous integration is reshaping how systems can be integrated and partitioned by technology. The trend is to exploit the height dimension by stacking die.

CONCLUSIONS

The rapid evolution of new threats is escalating the needs for more sophisticated radar, communications and EW platforms. Commercial applications like 5G and Satcom are sharing some of the same technical advances to fulfill their missions. Advances is digitization and signal processing enabling new architectures driving changes to the RF chains and the front-ends, which can no longer be considered in isolation. The RF portion, whether CMOS, SOI, SiGe, GaAs or GaN, must be optimized around the digitization and antenna-level requirements, and eventually co-integrated with the digitizer at the package level. While continuing to push operating frequency is still important, improving the power efficiency and facilitating co-packaging are just as critical in balance with dynamic range attributes such as output power, noise and distortion.

ACKNOWLEDGEMENTS

The author would like to thank colleagues Michael Neumann, Qui Luu, Marek Hempel, Fatih Kocer, Paedar Forbes, John Kilpatrick and Cliff King for discussions.