NETWORLD EUROPE

Strategic Research and Innovation Agenda Webinar
12,13th January 2023

DEVICES AND COMPONENTS

André Bourdoux (IMEC)

Table of content

25 pages
11 sections
>20 contributors
1 editor

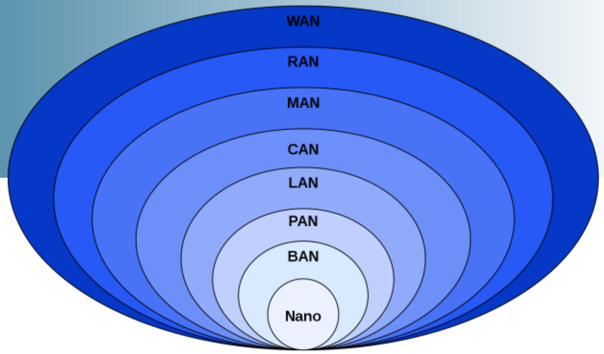
Vision and context

NETWORLD EUROPE

- Wired and wireless networks are in constant evolution
  - addressing all relevant societal challenges
  - support the digitization of the industry and society
  - support new applications
  - support the “more AI” trend
- Requirements on future networks
  - support very low to very high throughputs, increase area coverage, reduce latencies, improve reliabilities, integrate artificial intelligence, ...
  - support an ever-larger number of verticals.
- Antagonistic requirements
  - devices densities, AP/BS densities, (aggregate) throughputs are exploding
  - power
  - cost
- High impact on hardware, software and technologies

Sub-10GHz

NETWORLD EUROPE



https://commons.wikimedia.org/w/index.php?curid=95509836

- Standards trend: more and more standards!
  - Dominance of Cellular (xG), Wi-Fi, Bluetooth, GPS
  - Many other systems: NFC, IoT, WAN, NTN, …
- More efficient use of spectrum
  - XXXX-QAM, MIMO, multiple bands
- Higher integration
  - SoC (RF, analog, ADC/ADC, DSP, CPU)
- Front-end module and antenna nightmare
  - RF filters, switches, multi-band antennas, multiple channels, duplexers, …
- ADC/DAC "closer" to the antenna/RF
- Technology choice
  - Best RF (linearity, filtering), fastest digital (digital RF), lowest power
- New opportunities, new challenges
  - e-health: wearables, implantables, ingestibles, …
  - Human-machine interface, brain-machine interface
  - Mobile display technologies: AR/VR/XR, glasses, …
  - Sensing: passive, active (! self-interference)

- Better EVM
- Wider bandwidths
- Higher spectral efficiency
- Low cost, low power

This may look like "business as usual" but the requirements and constraints keep growing and so does the number of different platforms

Millimeter-wave and TeraHertz

- The 90 to 300GHz range has great potential
  - Fronthaul/backhaul, Access, P2P, sensing, ...
- At higher frequencies:
  - Higher free-space loss must be compensated by higher antenna gain
  - Front-end becomes more challenging
  - Ultra-wide bandwidth → multi-GHz baseband and 10+Gsps ADC/DACs
  - Improvements in circuit design needed
    - Phase noise, noise figure, IQ mismatch, ...
    - Frequency dependent effects: group delay distortion in all components
    - Efficient (hybrid) beamforming remains a challenge, especially for high gain, large bandwidth (phase shifters vs true -time delay, beam squinting)
  - Chip, chip interconnect and antenna module must be co-designed
    - Minimize interconnect lengths, losses
    - 2D, 2.5D and 3D electromagnetic simulations
- Air interface design exploiting ultra-wide bandwidth, very directional beamforming and “front-end friendliness”
- Move digital processing to analog (equalization, synchronization, ...)

$$P_r = \frac{P_T G_T G_R \lambda^2}{(4\pi R)^2} = \frac{P_T A_T A_R}{(\lambda R)^2}$$

Millimeter-wave and TeraHertz

- Solid-state technologies
  - Scale CMOS no longer the panacea: must be replaced or *complemented* with III-V
  - Huge trade-off:
    - Chip partitioning
    - Improve RF circuits vs calibration vs digital compensation
    - Technology choice
      - Silicon-based: RF-SOI, FD-SOI, SiGe BiCMOS
      - III-V on native substrates: InP
      - III-V co-integrated with CMOS
  - With further scaling, CMOS will transition from FinFET to gate-all-around structure (nanosheet, 2nm and beyond)
    - Impact on 10+Gsps ADC/ADC
- Not only wireless comm: convergence of communications and sensing
  - Passive THz imaging
    - Above-IC bolometer: better performance but expensive
    - Monolithic CMOS-based imagers: much lower performance but cost-effective
  - Active mm-wave and THz imaging
    - Higher frequencies enable smaller devices/better angular resolution and larger bandwidth/range resolution
    - Antenna options include on-chip and on-package
  - Imaging at >100GHz expected to boom and help driving circuit and technology research towards higher performances, smaller form factors and lower cost

Ultra-low Power Wireless

- IoT devices to grow to 100 billion by 2030
  - ULP sensors
  - e-health: wearables, implantables, ingestibles, brain-machine interface, active eye lenses, ...
- Huge challenge towards zero or near-zero power
  - Profound impact on the complete transceiver architecture and design and the protocol
- (nearly) Battery-free operation
  - Energy scavenging (thermoelectric, photovoltaic, piezoelectric, RF or wireless, wind, vibration, ...)
  - Wake-up receivers
  - Back-scattering devices
- Degradable devices
  - Huge # devices → huge e-waste at end of life
  - Bio-degradable substrates
  - Renewable materials

Antenna and Packages

Moving to >100GHz brings new challenges

- Packaging for consumer equipment is a challenge
- Lossy interconnects → avoid interconnects → on-chip or in-package antenna
- Lenses can help to increase the gain but reduce the field-of-view
- Dual polarization (because of mobility)
- Wide bandwidth
  - Frequency-dependent behaviour
- Mismatches and angle-dependent coupling between antenna elements
  - Need for efficient array calibration
  - Can be expensive at manufacturing time
  - Fully-automatic, both start-time and run-time
- MIMO arrays, hybrid beamforming architectures
- New phased array architectures and interconnects
  - Substrate integrated waveguide

Antenna and Packages

Metamaterials and metasurfaces

- Allow to manipulate electromagnetic waves
- → huge potential for antennas and surfaces ("Intelligent Reflective Surfaces")
- Coverage increase, smart radio environment, better cell edge coverage, less interference and electromagnetic pollution
- Can be combined with antenna technologies, massive MIMO, mm-wave and THz communication, D2D
- Applications include:
  - Antenna design, absorbers, reflectors, super-lenses, cloaking devices, RCS manipulation (radar), ...
- Very active field of research, many innovations and disruptions

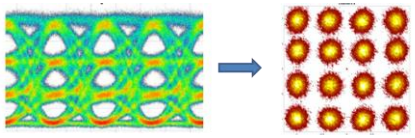
High-speed Transceivers, Wireline and Optical

Radio-over-fibre communication, sub-systems and components for B5G and 6G networks

- Fronthauling needs explode with coordinated BF, CoMP, massive MIMO and cell-free MIMO
- CPRI and OBSAI are expected to saturate
- Example:
  - 2GHz BW, 4 carriers, 3 sectors each with 32 antennas, 8bits I&Q, 8B/10B encoding, 10% overhead → sustained throughput of 25Tb/s
- Innovative fronthauling solutions are needed
  - Analog RoF (high linearity needed)
  - RF Sigma-delta modulation
  - ...
- Split processing trade-offs between RRH and BBU

Towards Terabaud capable opto-electronic transceivers

- Traffic to Data Centres is exploding
  - Higher telecom needs + cloud-based ML/AI + ...
- Need for new generations of optical transceivers with ever higher capacity
- Deployment of optical links at ever shorter distances
- More pervasive use of coherent transceiver technologies
  - From long-haul to metro to data centers to access
- Need for electro-photonic Systems-in-Package and co-packaged optics
  - Optical transceiver chiplets + CMOS data processing in one package
- Increase
  - Symbol rate: 100G → 200G → 400G → 800G → 1.6T → 3.2T ... transceivers
  - Number of parallel lanes: (multiple wavelengths and/or fibres)
  - Higher spectral efficiencies: 4-PAM → complex modulation
  - Integration: denser integration e.g. 3D modules
- Enabling technologies:
  - CMOS → SiGe, InP
  - Novel materials for ultra-broadband optical modulators and detectors: e.g Organic hybrid material, Ferro-electric materials, Lithium Niobate (LiNbO3)
  - Monolithically integrated optics and electronics
  - Optically assisted analog-to-digital and digital-to-analog conversion



High-speed Transceivers, Wireline and Optical

Ultra low-cost and low-power coherent "lite" transceivers

- Need for coherent detection for shorter ranges and at very low-cost
- Potential enabling technologies:
  - Integrated narrow linewidth laser sources
  - Integrated optical phase locked loops
    - For carrier recovery
  - Novel equalization approaches relying on co-developed opto-electronics
    - Move compute-intensive digital functions to optical e.g. passive optical filter

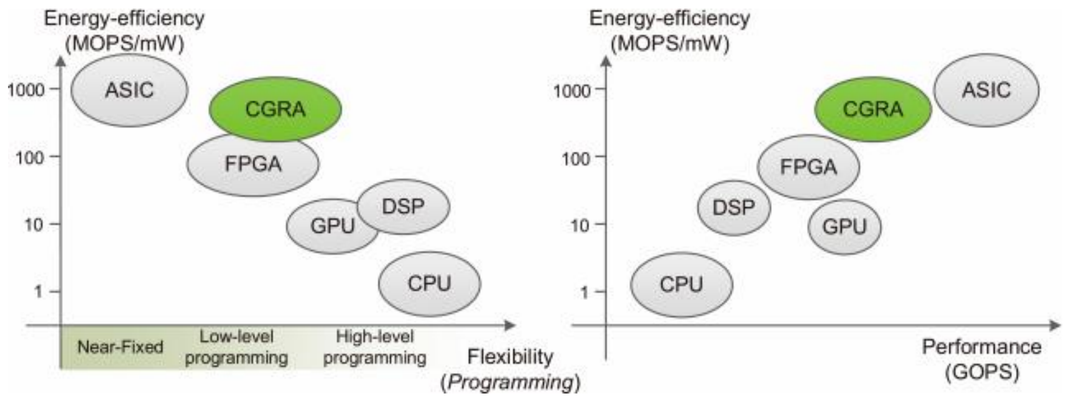
Optically assisted wireless subsystems

- Microwave photonic techniques to replace conventional beamforming
- …

Baseband Modems

- Architectural trade-off
  - Flexibility/Programmability
  - Performance
- According to application
  - IoT, UE, Infrastructure
  - SISD, SIMD, MIMD
- Trends
  - Coarse-grained reconfigurable architecture
    - Near-ASIC performance with SW-like programmability
  - Array of COTS CPUs
    - Better suited for infrastructure
  - Deep-learning architectures for PHY processing
- Challenge for very high throughput: from nJ/bit to pJ/bit
- Challenge for very low throughput: from sub-mW to sub-µW
- Memory architecture/hierarchy!
  - A lot of power goes in data exchange

Picture from : [“A Survey of Coarse-Grained Reconfigurable Architecture and Design: Taxonomy, Challenges, and Applications”, Liu et al., 2019]



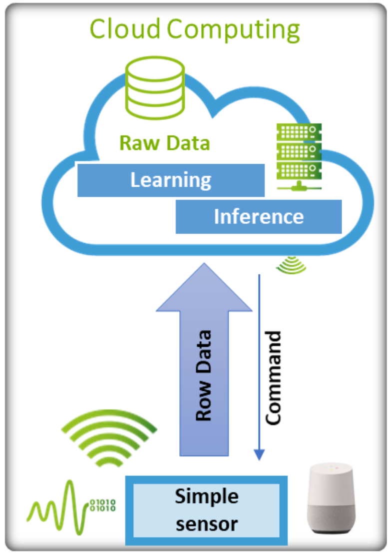
CMOS scaling

- Towards 1nm (A10) and beyond (A7, A5, ...)
- Lithography evolution
  - Extreme UV
  - Multi-patterning EUV
  - High numerical aperture UV
- Transistor architecture evolution
  - FinFET
  - Gate all-around (nanosheet, forksheet)
  - Complementary FET
- Innovative interconnect architectures
- 2.5D and 3D stacking

Picture from : https://www.imec-int.com/en/articles/20-year-roadmap-tearing-down-walls

- “Tearing down the walls”: the power wall, the cost wall, the memory wall

Processors for Cloud-AI, Edge-AI and on-device-AI



Past
Present
Future

AlexNet

Cloud computing is still the workhorse today

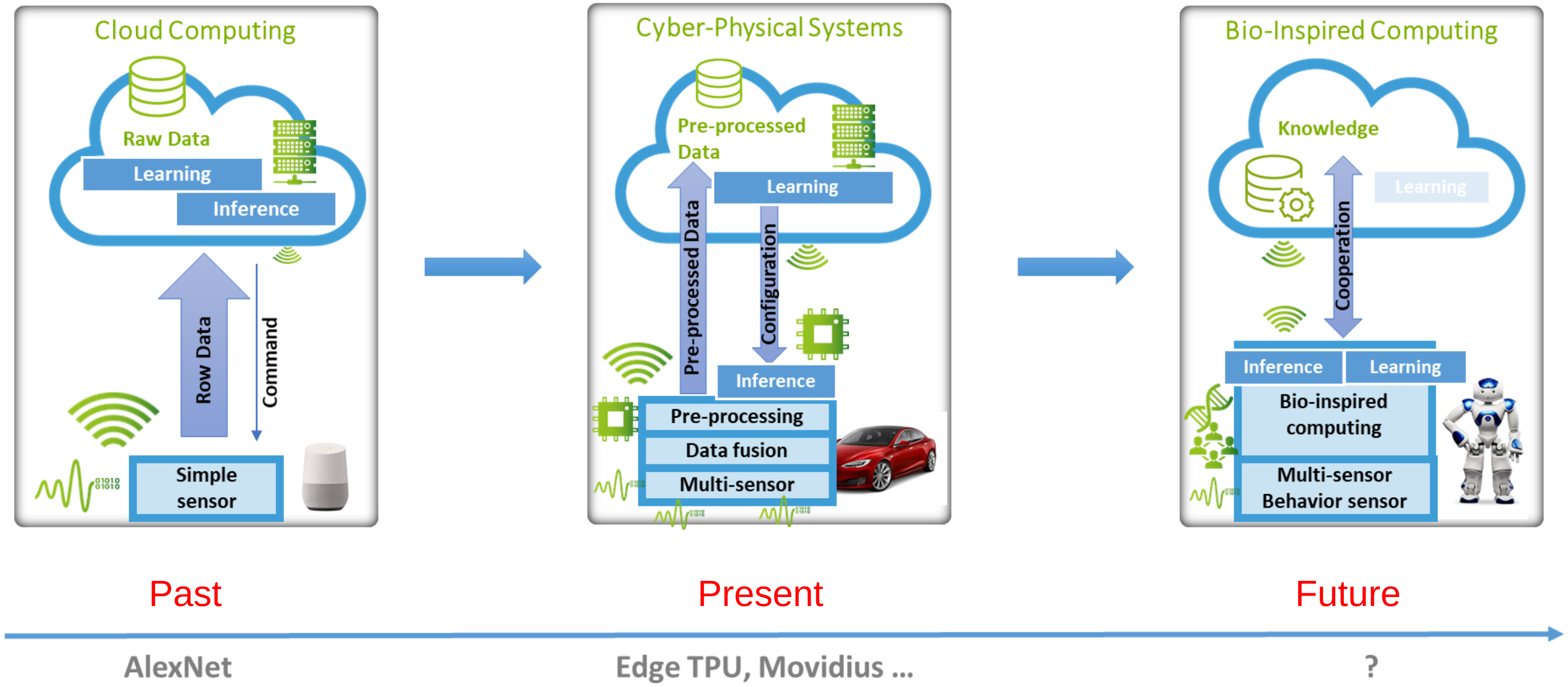
Need for very **high compute power for training** large networks

→ Large GPUs, TPUs, scalability over multiple nodes

Need for **low latency inference** (on a batch of 1 piece of data)

→ FPGA, CPU

Processors for Cloud-AI, Edge-AI and on-device-AI



But there are ever larger needs for edge computing

Because of safety of operation, latency, privacy, power dissipation …

→ Need for dedicated ASICs, with sensor integration

Processors for Cloud-AI, Edge-AI and on-device-AI

Several techniques can be employed at architecture and circuit levels

Increased computing efficiency

Weight quantization

Reduced bit accuracy
- Smaller memory footprint
- Lighter operations

Variable bit precision

Handling higher bit accuracy when needed
- For higher inference precision

Sparsity

Skip MAC operations
- When weight or intermediate result is 0

Increased storage efficiency

Near memory computing

Avoid external memory accesses

Weights
- Embedded Non-Volatile Memory

Intermediate results
- SRAM or Embedded DRAM

In-Memory computing

SRAM or Embedded NVM

Digital or analog

A completely different paradigm, related to CIM:

Spiking neural networks (SNN), analog and digital flavour

Memories - Memories for processing and storage

NETWORLD EUROPE

We are entering the zettabyte and soon the yottabyte eras: yearly growth rate: 1.2 ... 1.4x

- Yottabyte predicted in 2030
- Data and traffic generated through
  - Apps such as Amazon, YouTube, Facebook, Netflix ...
  - IoT such as autonomous cars, smart buildings, smart city, e-health, ...
- Huge environmental problem: heat and power consumption
- 3-axis performance improvements
  - Density
  - Time: latency and speed/bandwidth
  - Energy
  - ... and Cost ...

Memories - Memories for processing and storage

	ENERGY REDUCTION FOR A GIVEN THROUGHPUT	DENSITY IMPROVEMENT	SPEED (AT DEVICE LEVEL)
CACHE (SRAM)	1.12x	1.15x	1.1x
MEMORY (DRAM)	1.1X Cs reduction, Vdd reduction	1.2X ➔ 1.1X Slow down with C scaling	1x
STORAGE (FLASH)	1X	1.4X ➔ 1.2X 3D log trend cannot last	1x performance increase from 2D to 3D
ARCHIVAL (TAPES)	1.1x	1.4X Doubling at each LTO node	1.1x

https://www.imec-int.com/en/imec-magazine/imec-magazine-september-2018/emerging-memories-for-the-zettabyte-era

- Growth rate for SRAM and DRAM is saturating and not sufficient towards yottabyte era
- Novel approaches and technologies are needed to sustain the growth rate
  - 3D stacking, already largely exploited
  - Emerging storage class memories to fill the gap between DRAM and NAND
    - PC-RAM, VMCO, CB-RAM, OxRAM, ...
  - Magnetoresistive RAMs: many variants
    - STT-MRAM (for L3 cache), SOT-MRAM (for L1-L2 cache), VCMA-based MRAM, ...
- DNA storage
  - Highest density *potential* by orders of magnitude
  - Challenges: speed, reliability ... and cost
  - Very long term

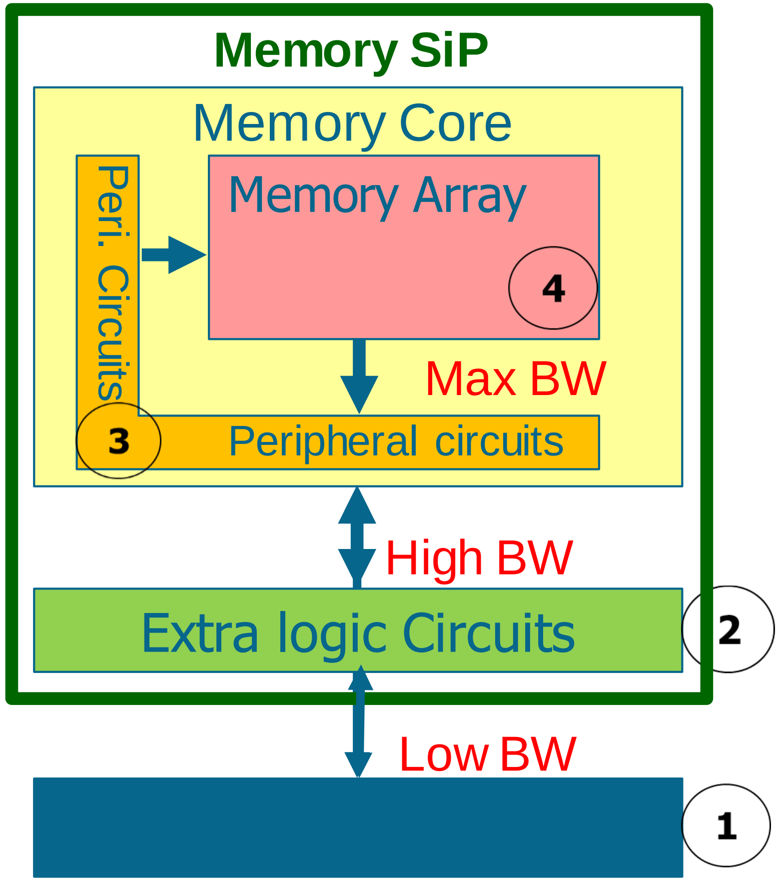
Memories - Computation-in-Memory (CIM)

Computer architectures: Classification

- Computation-outside-Memory (COM)
  1. Far
  2. Near
- Computation-in-Memory (CIM)
  3. Periphery
  4. Array

CIM relies on memristive devices

Not a mature technology



Memories - Computation-in-Memory (CIM)

NETWORLD EUROPE

Challenges

Technology

- Multi-state behaviour
- Energy switching
- Threshold behaviours
- Endurance
- Fault tolerance
- Variability, ROFF/RON ratio
- Integration, yield

Circuit design

- High precision programming of NVM
- Fast and energy efficient signal conversion circuits (DAC, ADC)
- Precise measurement of current
- Vector x matrix: output as current
- Control complexity

Architecture

- Micro vs macro architectures
- Intra- and inter-communication
- System accuracy
- Design exploration
- Simulation tools

Tools & applications

- Mapping applications on architecture
- Compilers
- EDA tool chains
- Bridging device characteristics to circuit and to algorithm design
- Simulators

Research on Sustainable Security and Privacy - Motivation

Today's Limitations:

- Devices with limited security life-time
- Devices do not survive attacks – require manual recovery
- Manual mitigation of risks and frequent patching
- Crypto is degrading and suffers against quantum computing attacks

Desired Future:

- Devices that survive for years in the field
- With minimal maintenance and automated recovery
- Guaranteed long-term survival of crypto mechanisms

NETWORLD EUROPE

Research on Sustainable Security and Privacy - Research Vectors

Higher Reliability | **Tolerating Faults & Attacks**

Defend | Contain | Detect | Recover | Repair

Research:

Maintaining Security and Surviving Attacks

- Graceful degradation into fail-secure states, maintaining critical services
- Systems survive attacks with automated recovery

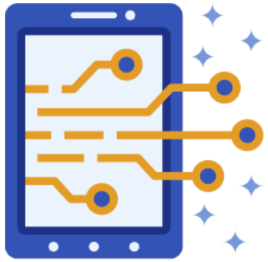
Research:

Post-Quantum Cryptography with Hardware Support

- Range of crypto that are robust against quantum computing attacks
- Toolboxes for wide range of usages

Opportunities for IoT Components and Devices

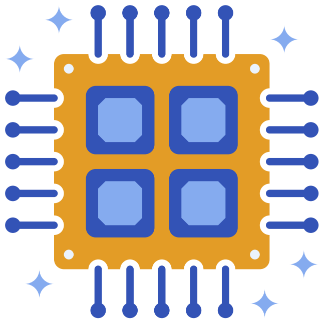
IoT - Components and Devices – Research areas



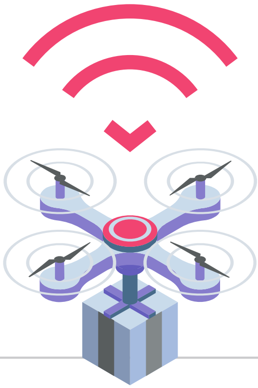
- **Pervasive wireless connectivity** as a major component behind the IoT technology and one of the key layers in IoT and IIoT architecture.
- Research challenges in the development of IoT components and devices for IT/OT integration using **multi-frequency/multi-protocol heterogeneous wireless communication and networking** for IoT/IIoT and edge computing with built-in end-to-end distributed security.
- **Ultra-low power IoT,** extended to Tactile IoT components and on-IoT device AI techniques and methods.
- **Wide frequency range** from sub-1GHz to THz
  - Use of CMOS and III-V semiconductors-based GaAs, GaN, InGaAs, SiC semiconductor technologies. Integrate microwave and analogue front-end technology and millimetre wave monolithic integrated circuits (MMIC).
  - Requires alignment between SNS and KDT.

Opportunities for IoT Components and Devices

Approach for IoT devices



- Specialized IoT devices and sensors enabled and validated especially for vertical sectors
  - Leveraging system on chip activities.
  - Specifying the way to communicate in the network/systems .
  - Integrating them in their operational systems in vertical (and as well cross- vertical) application domains.
- Sustainable growth for energy efficient IoT devices development, battery efficiency and battery-free operation.
- Degradable devices and energy autonomous devices that uses ultra-low power radios and harvest the needed energy.



SRIA 2022 Chapter 9 Contributors

- André Bourdoux (editor)
- Jan Craninckx
- Piet Wambacq
- Aarno Pärssinen
- Didier Belot
- Yao-Hong Liu
- Ullrich Pfeiffer
- Antonio Manzalini
- Peter Ossieur
- Fredrik Dahlgren
- Achouche Mohand
- Maziar Nekovee
- Alexandre Valentian
- Said Hamdioui
- Francky Catthoor
- Gouri Sankar Kar
- Valerio Frascolla
- Ovidiu Vermesan
- Luis Perez-Freire
- Georgios Karagiannis
- Damir Filipovic

public

NETWORLD EUROPE

THANK YOU FOR YOUR ATTENTION