

# ***Embedded Reconfigurable System Integration***

***Decentral Adaptivity, Reliability and Intelligence as Enablers***

**Jürgen Becker**

**Karlsruhe Institute of Technology – KIT**

Institute for Information Processing Technologies (ITIV)

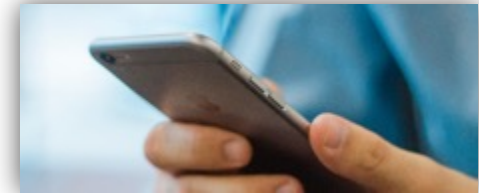
ITIV



# Context and Motivation

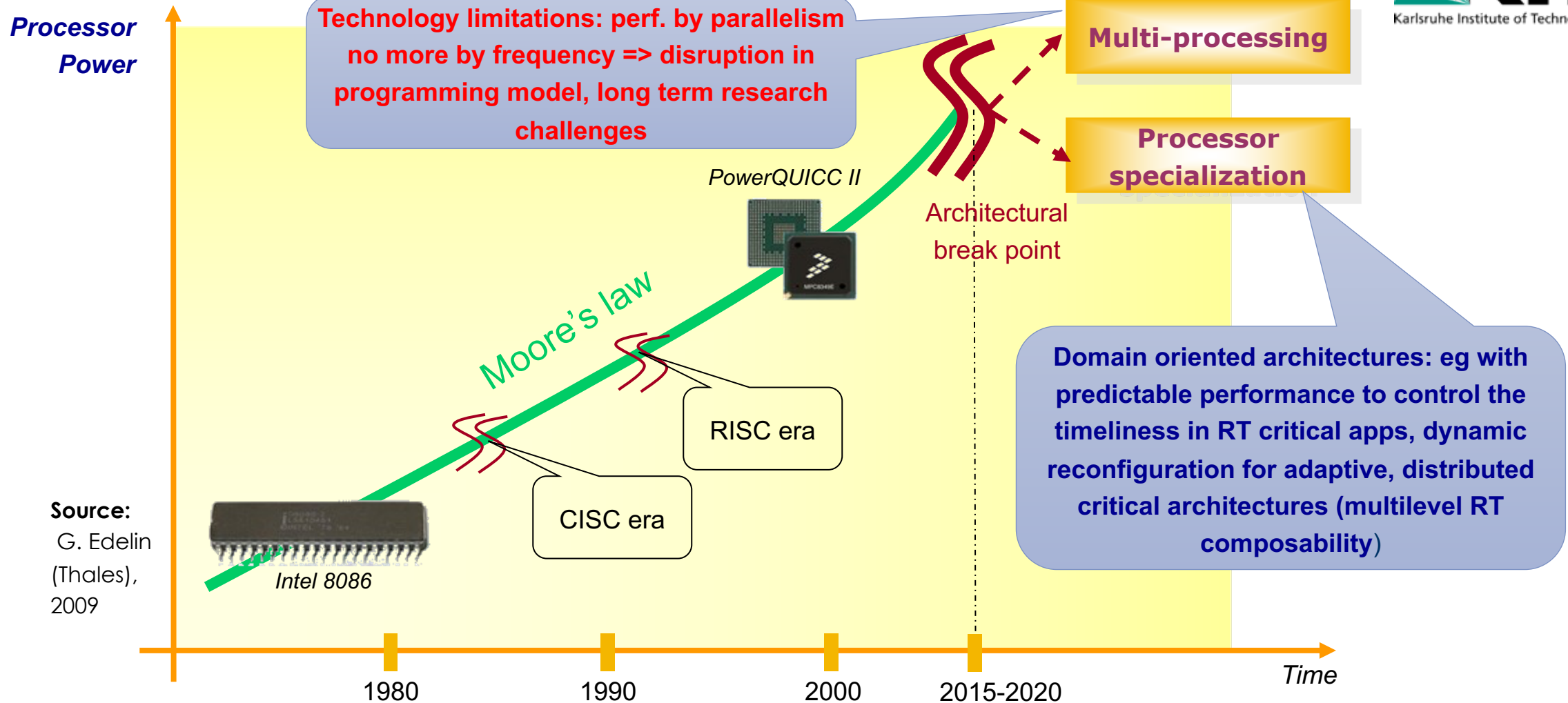
## ■ Current embedded systems are subject to challenging requirements:

- Increased **performance** is necessary to facilitate the execution of computationally intensive algorithms (machine learning, real-time control, big data, ...)
- **Power and energy consumption** must be minimized to facilitate the constraints of mobile and wireless devices
- A sufficient degree of **dependability** is necessary to employ digital systems in safety-critical environments (autonomous driving, ...)



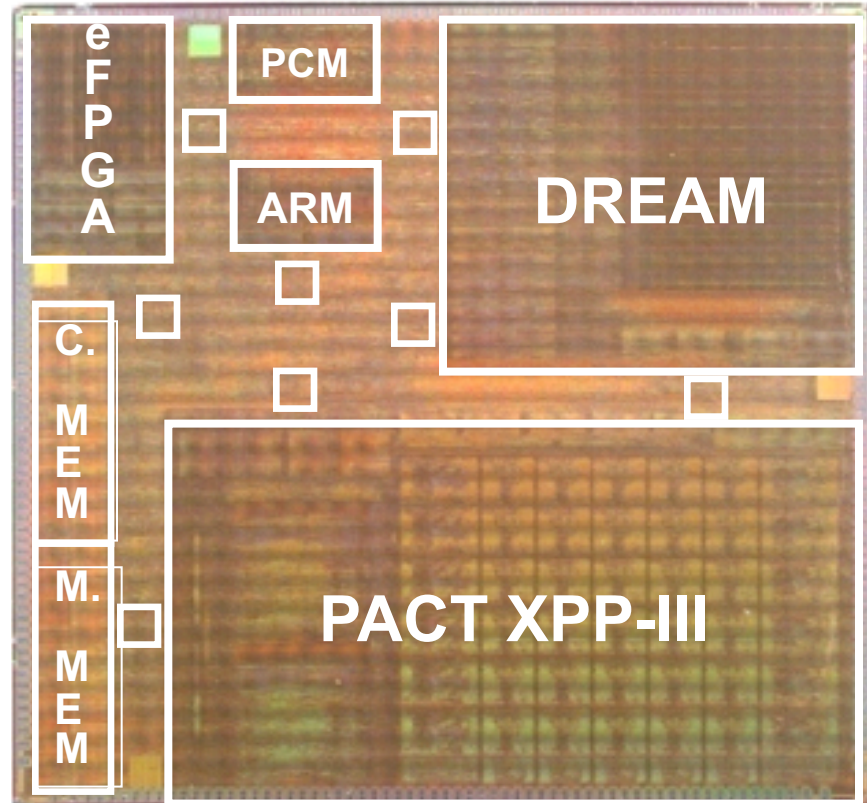
- Existing technology must evolve in order to meet these requirements  
⇒ **Adaptivity, HPC, AI and Reliability** can play key roles in this research process

# Motivation: Processor Developments ....



**Major architecture disruption: multiprocessing & specialization have strong impact on software**

# Former EU-Project Morpheus: Chip Integration Results



Technology: STM-CMOS090GP  
Supply voltage: [.9V : 1.2V], ref 1V  
Transistor count: 97 M  
Chip area: 110 mm<sup>2</sup>  
Static power: 235 mW  
Dynamic power:  
Standby: 700 mW  
peak: 3100 mW  
Pinout: 256, 163 I/O

Main (System) DOMAIN:  
Frequency@1V: 250MHz  
Dynamic power: 2.8 mW/MHz

Source: G. Edelin, Thales, 2009

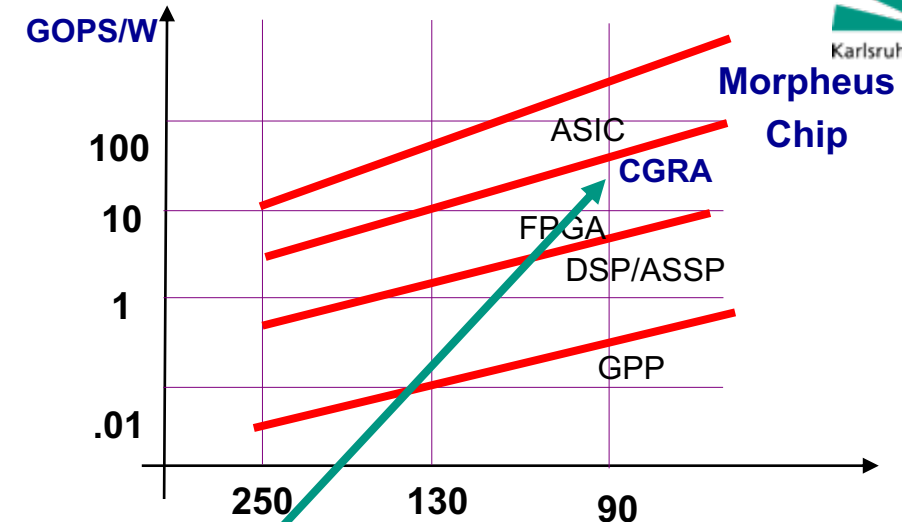
<b>XPP subsystem:</b>	<b>DREAM subsystem:</b>	<b>eFPGA subsystem:</b>
Macro area: 8.5x5 mm <sup>2</sup>	Macro area: 5x4.5 mm <sup>2</sup>	Macro area: 1.8x2.8 mm <sup>2</sup>
Max Freq@1V: 150 MHz	Max Freq@1V: 200 MHz	Max Freq@1V: 100 MHz
Dynamic power: 13 mW/MHz	Dynamic power: 1.7 mW/MHz	Dynamic power: 1.3 mW/MHz



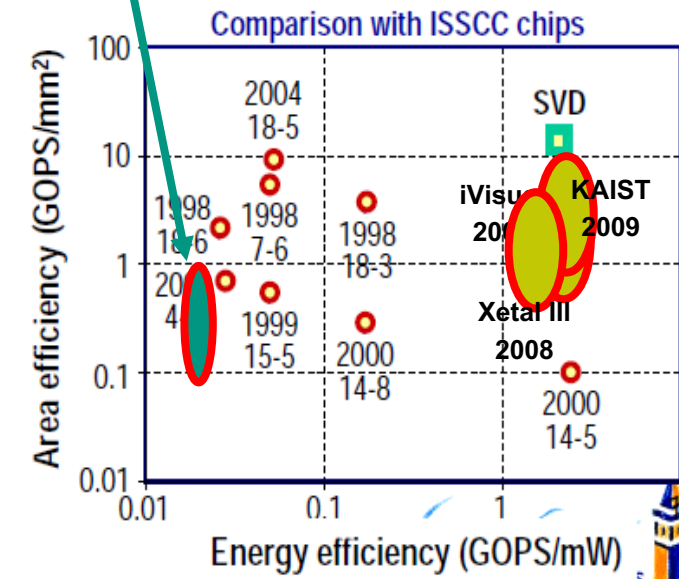
# Morpheus Results: Performance Analysis Comparison

Device	Techno	GOPS	GOPS /mm <sup>2</sup>	GOPS/ W
<b>Morpheus</b>	<b>CMOS090</b>	<b>60</b>	<b>0.6</b>	<b>20</b>
ARM9	CMOS090	0.35	0.15	1
ST Asic	CMOS090	n.a.	10	1000
TU iVisual (JSSC 2008)	CMOS180	77	1.16	205
KAIST Vect Proc (JSSC 2009)	CMOS130	125	3.4	214
Philips Xetal II (JSSC 2009)	CMOS090	107	1.44	170
Cell		200	n.a.	1:2
UCB Bee2	Fpga	n.a.	n.a.	2.2
Xilinx Virtex-II 8000	Fpga	450	n.a.	3.5

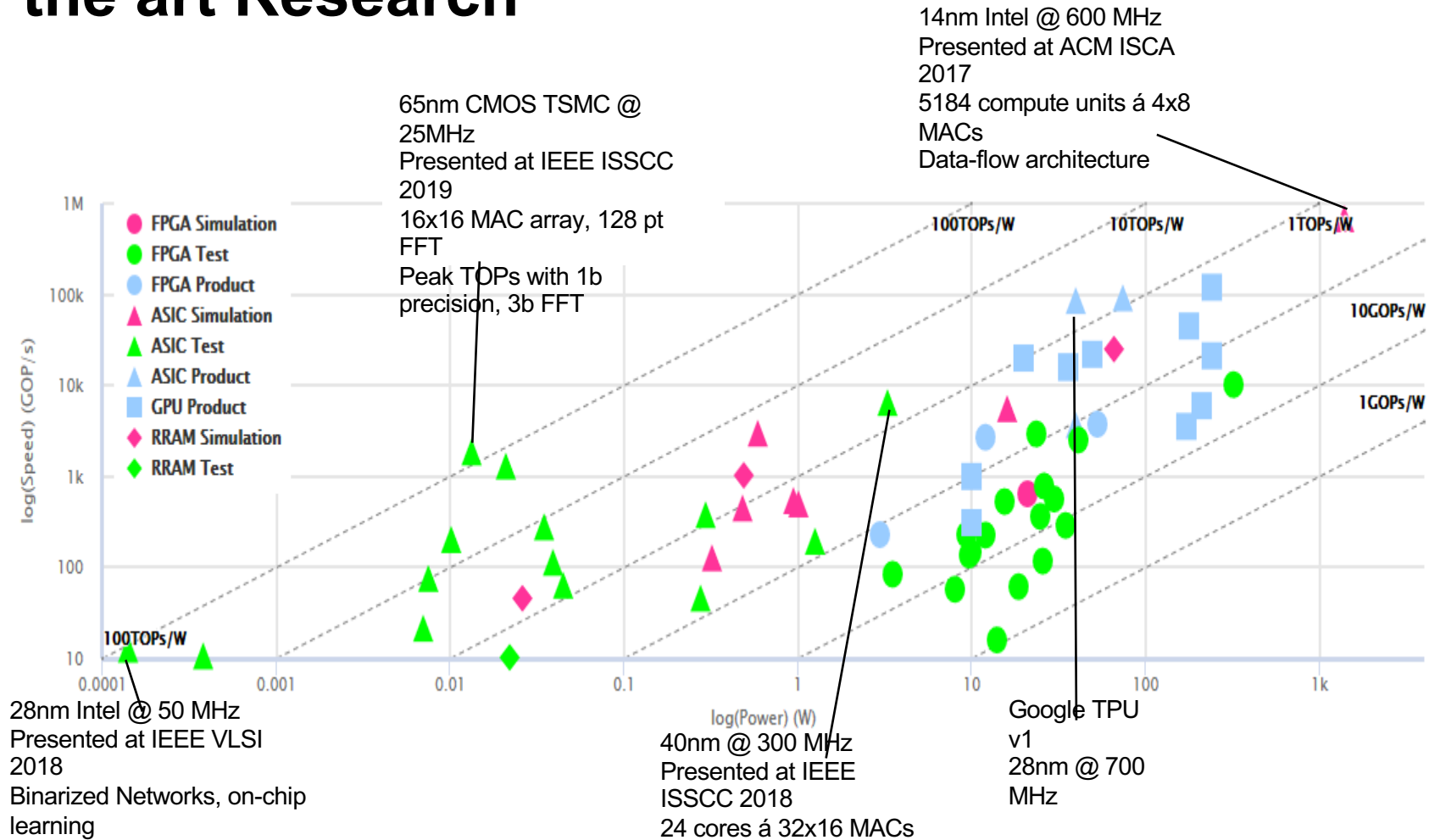
Source: G. Edelin, Thales, 2009



Source: T.Claasen (ISSCC99)



# State of the art Research



## Comparison of state of the art publications on CNN accelerators

- Power in Watt vs. „Speed“ in GOP/s

Source: <https://nicsefc.ee.tsinghua.edu.cn/projects/neural-network-accelerator/>

# State of the art Chips ...

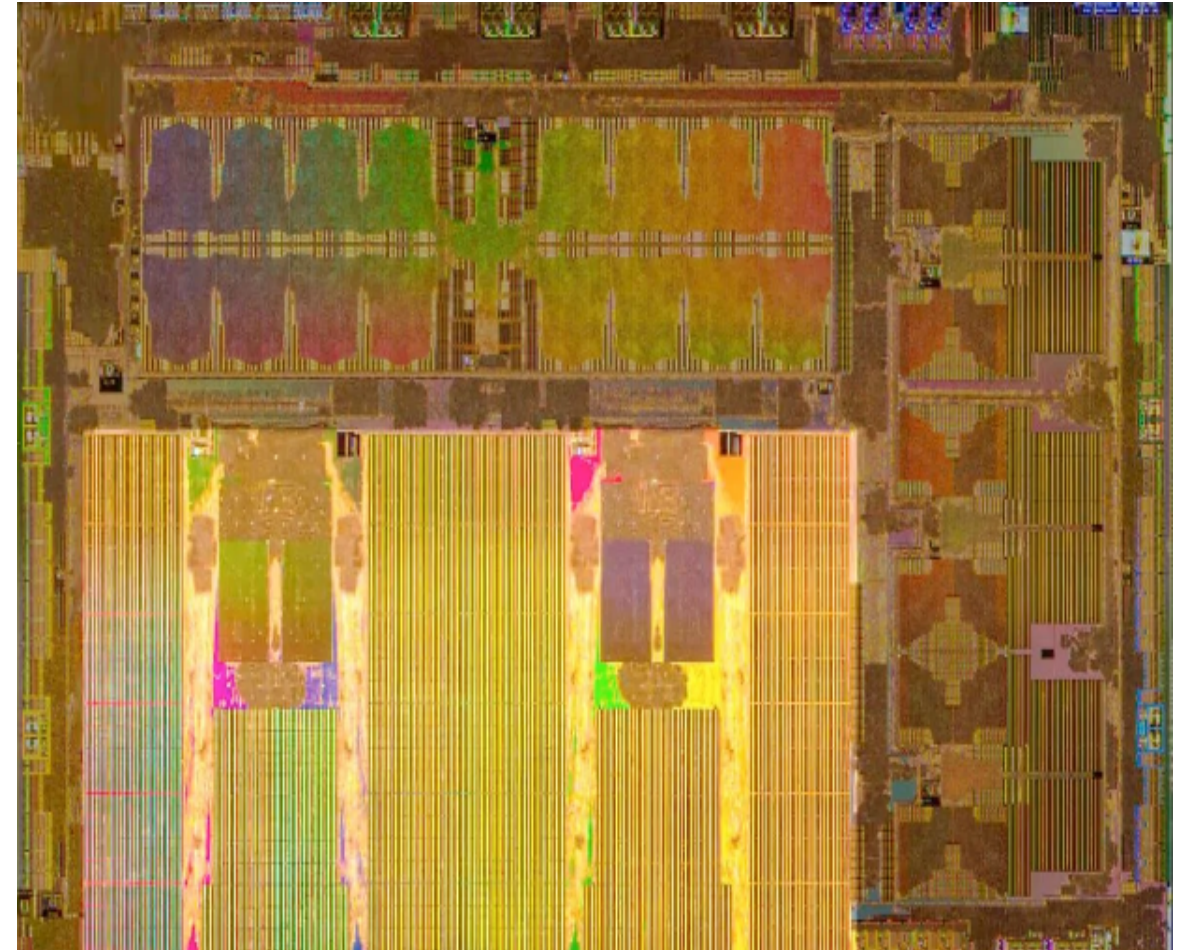
## Tesla – Full Self Drive Chip

- Six ARM CPUs
- Dedicated GPU

- **Two NPUs**

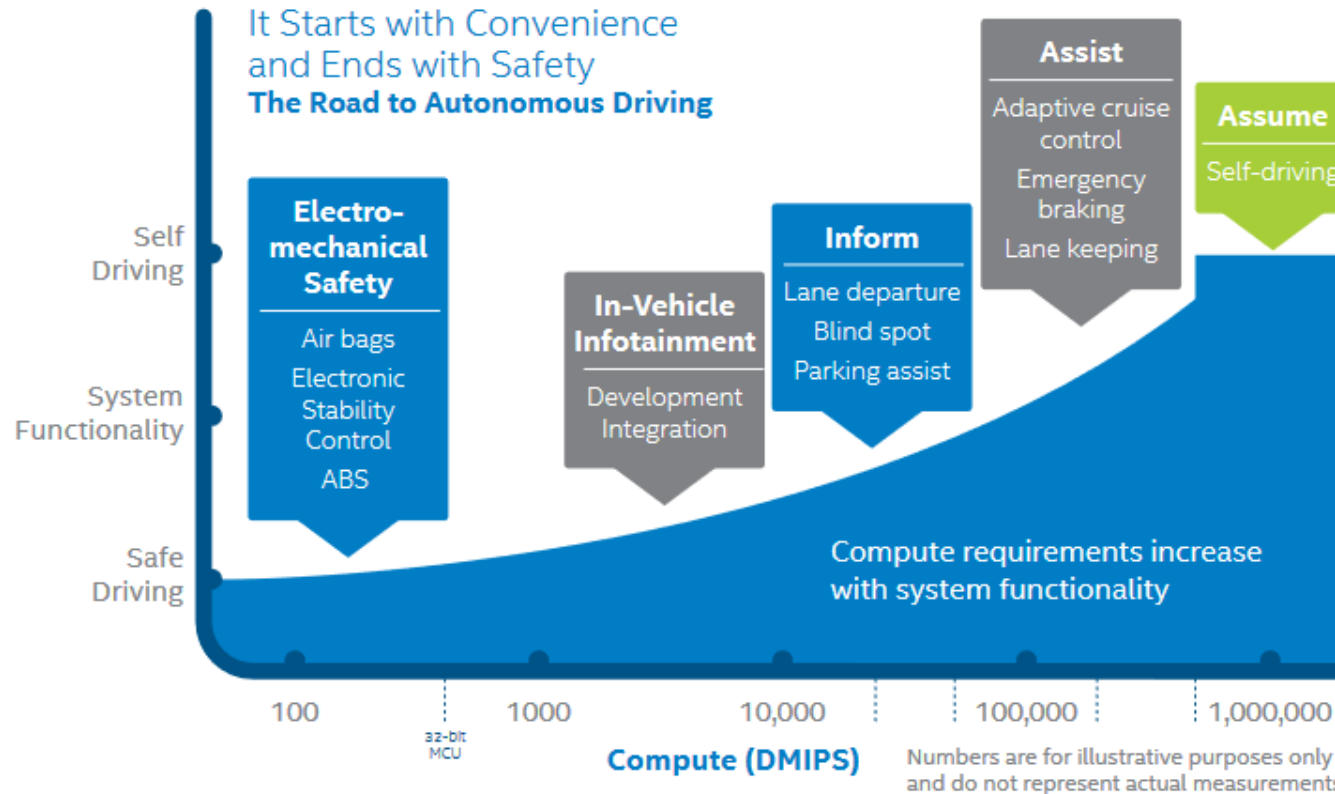
*Neural Processing Unit (NPU)*

- *Systolic Array*  
with Memory
- 36 TOPs @ 32 W



# Automotive Technology Developments

## -> Autonomous Cars: Embedded HPC Demand



## Common carbon footprint benchmarks

in lbs of CO2 equivalent

Source: MIT Technology Review

Roundtrip flight b/w NY and SF (1 passenger)

1,984

Human life (avg. 1 year)

11,023

American life (avg. 1 year)

36,156

US car including fuel (avg. 1 lifetime)

126,000

Transformer (213M parameters) w/ neural architecture search

626,155



(1) <https://www.intel.com/content/www/us/en/automotive/driving-safety-advanced-driver-assistance-systems-self-driving-technology-paper.html>



# EPI – European Processor Initiative

- Project timespan: Dec. 2018 to End of 2021

- Target markets:

- HPC and Automotive

- Proposal drivers:

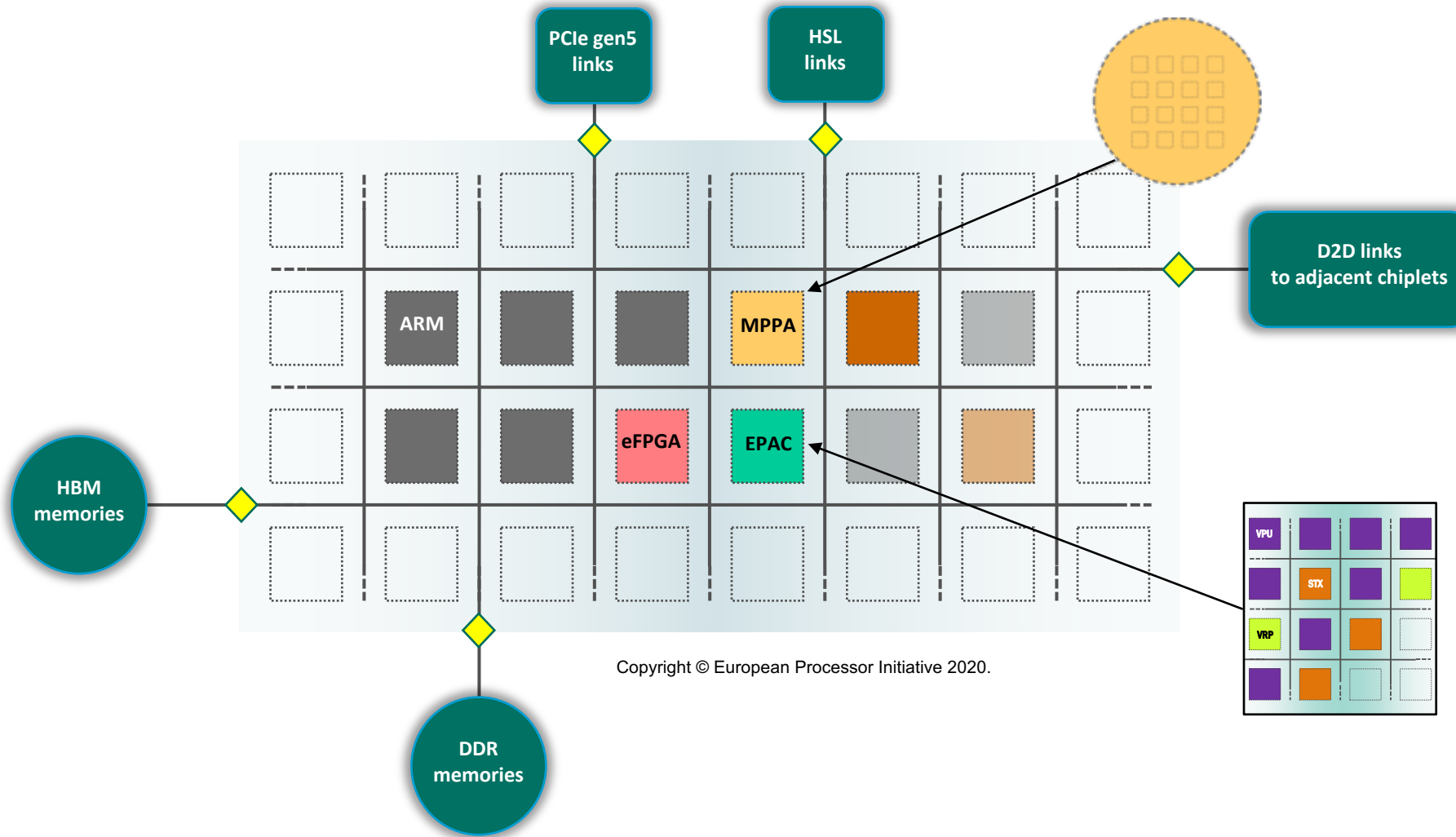
- Create a competitive European HPC and Automotive platform



→ **Mission:** independable EU Exascale machine by 2023

More Info: <https://european-processor-initiative.eu>

# EPI – Common Architecture



- **MPPA:**  
Multi-Purpose  
Processing Array
- **eFPGA:**  
embedded FPGA
- **EPAC:**  
EPI Accelerator

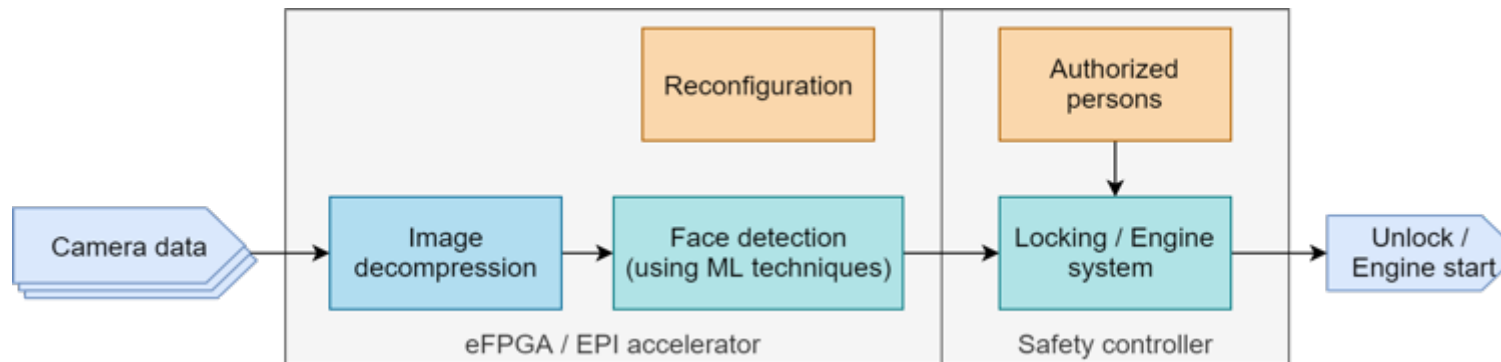
# EPI – eFPGA Integration

- **Provided by Menta S.A.S.**
- Optimized for general purpose and automotive applications
- **Application scenario in EPI: Run-time reconfigurable crypto and general purpose accelerators in a low power domain**

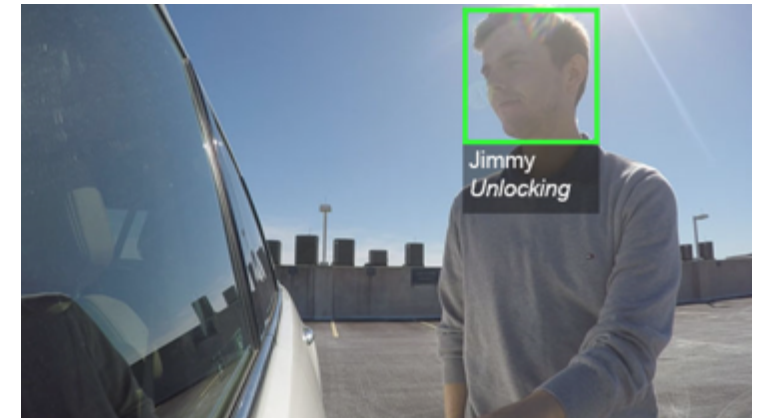


# EPI – eFPGA NN Application Scenario

- **Accelerator** for image processing with low-latency *Convolutional Neural Networks (CNN)*
  - Face detection to unlock the car when it is standing
- **Reconfiguration** allows for fast switching between different models for different applications



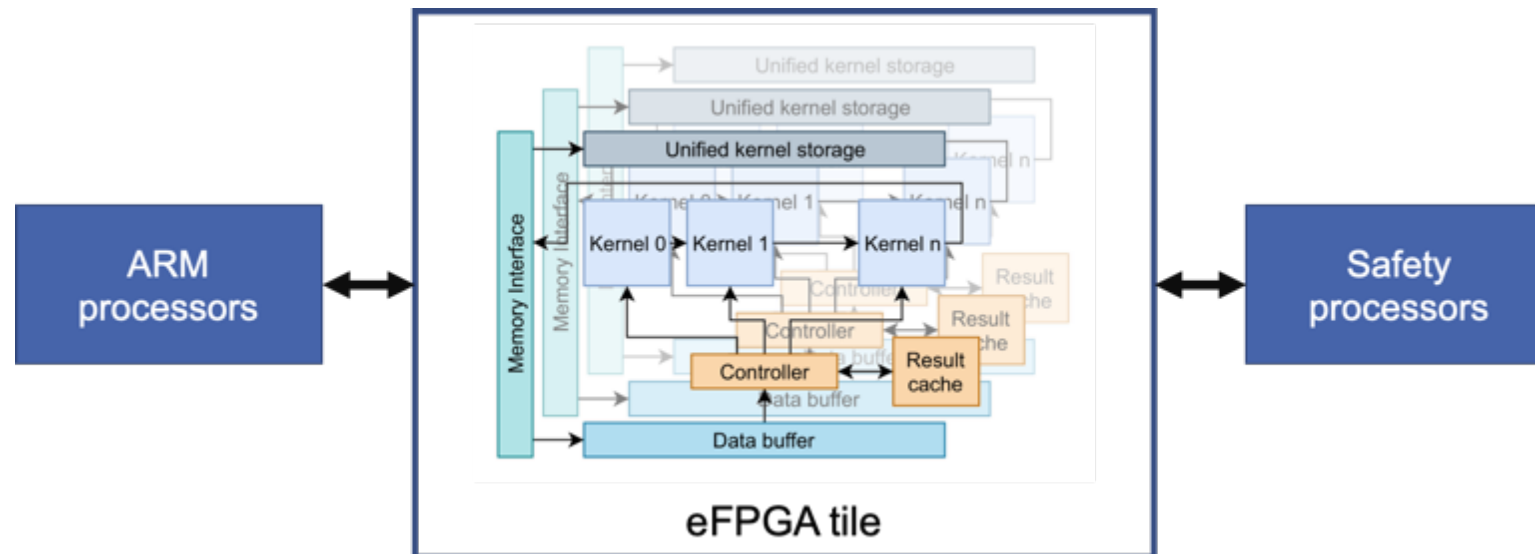
**System Overview**





# EPI – *eFPGA* current challenges and results

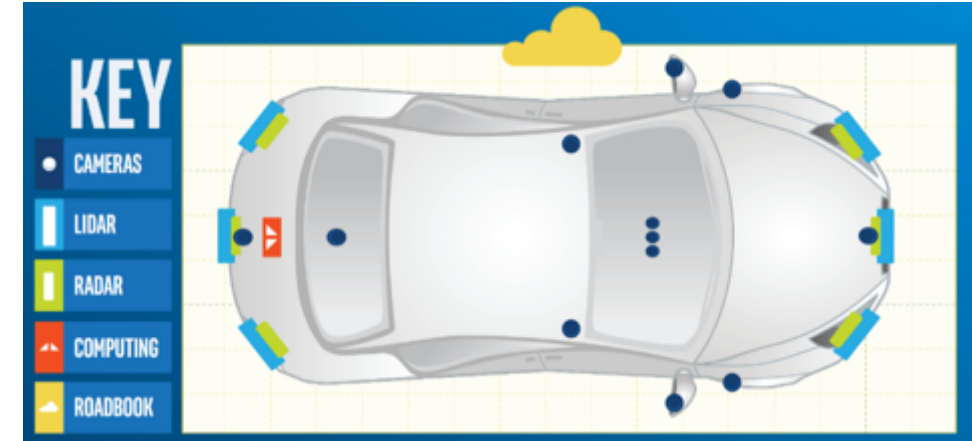
- Achieve a maximum performance with constrained eFPGA area (3mm<sup>2</sup> in 7nm)
- Design Space Exploration with a scalable CNN Accelerator
  - On 3mm<sup>2</sup> we can compute 52 convolution operations in parallel at 160 MHz
  - With about 70% overall utilization
  - The eFPGA is still completely reconfigurable and suits any other architecture



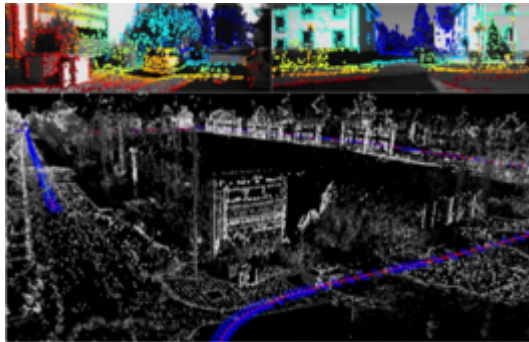
# Motivation – eHPC in Automotive *Innovation Driver* -> Autonomous Driving

## Performance Challenges in different Domains

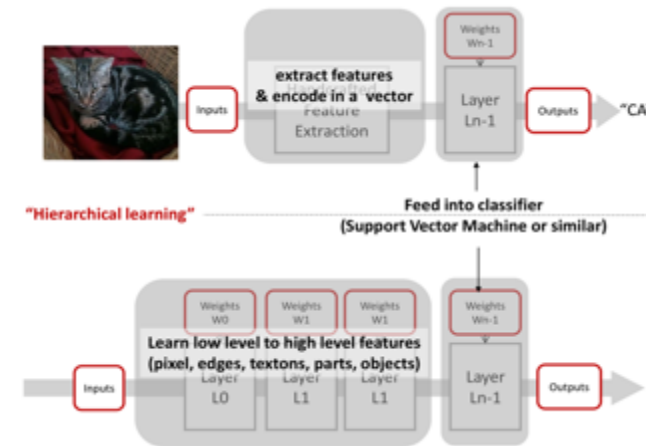
- **Computer Vision**
  - Recognition (e.g. Intel i7)
  - Image Classification (Xilinx Everest)
  - Semantic Segmentation (NVIDIA PX2)
- **Data fusion**
  - Cameras
  - Lidar, Radar
- **Connectivity / 5G (Intel Go)**



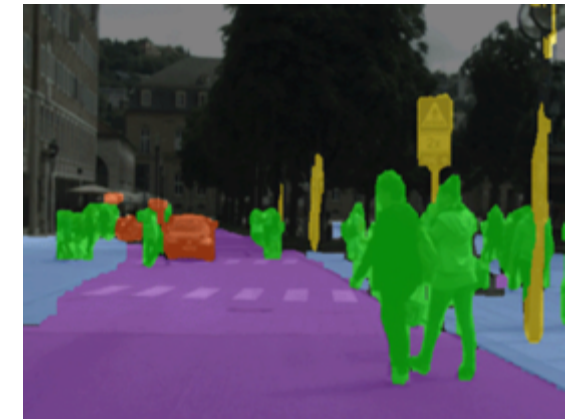
Taken from (1).



Taken from (2).



Taken from (3).



Taken from (4).

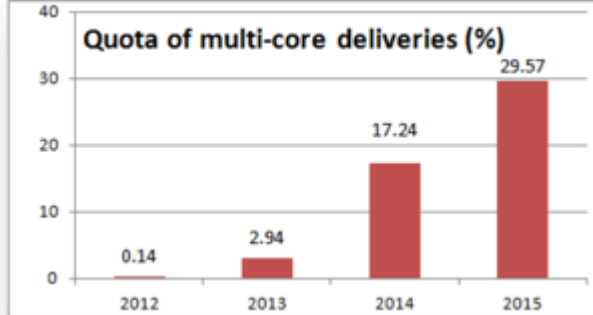
(1) <https://newsroom.intel.de/news/sensors-the-eyes-and-ears-of-autonomous-vehicles/>

(2) <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8237683>

(3) Xilinx, Inc. **Architectures for Accelerating Deep Neural Networks.**

(4) <https://blogs.nvidia.com/blog/2016/01/05/eyes-on-the-road-how-autonomous-cars-understand-what-theyre-seeing/>

# Embedded Computing Performance Needs



[Source: Shared SW development in multi-core automotive context, L. Michel, et. al, 2016]

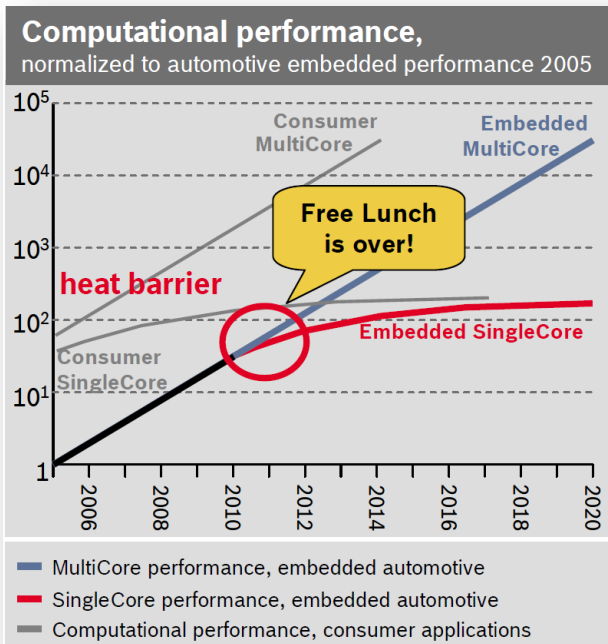
Quota of deliveries based on multi-core CPU at VW/AUDI (not yet in safety critical applications):



[Source: Audi AG]



[Source: Airbus Group]



[Source: The Challenge of Mastering Parallelism in Real-Time Systems, J. Haerdlein, 2014]

## Specified Requirements about ...

- ... **functional Safety** (*Safety*) ...
- ... **Data Security** (*Security*) ...
- ... „**Space, Weight & Power**“ (SWaP) ...

... dedicated and new Solutions needed in the Context of Multi-Core & MP Systems

**Safety/Security Codesign** is one new Challenge

-> additional **Requirements** for Platform Development!

# National Flagship Project *ARAMiS*

**Coordinator:** Karlsruhe Institute of Technology

**Duration:** 3 years **Start:** 12/2011, **End:** 03/2015

**Funded by:** German BMBF (Federal Ministry of Education and Research)

**Follow-up Project:** ARAMiS II, also by German BMBF

**Duration:** 3 years **Start:** 2016, **End:** 2019

GEFÖRDERT VOM



Bundesministerium  
für Bildung  
und Forschung



**OEM**



**Tier 1**



Audi  
Electronics Venture GmbH



**Semiconductor manufacturers**



For the support in  
certification issues  
regarding the rail domain

**Research**



TECHNISCHE UNIVERSITÄT  
CAROLO-WILHELMINA  
ZU BRAUNSCHWEIG



Fraunhofer  
IESE AiSEC



UNIVERSITÄT PADERBORN  
Die Universität der Informationsgesellschaft



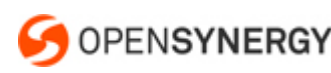
TECHNISCHE UNIVERSITÄT  
KAISERSLAUTERN

Universität  
Stuttgart

**Tool/software manufacturers**



WIND RIVER



SYMTA VISION



Deutsches Zentrum  
für Luft- und Raumfahrt e.V.  
Projekträger im DLR





# ARAMiS: Multicore Challenge Virtualization

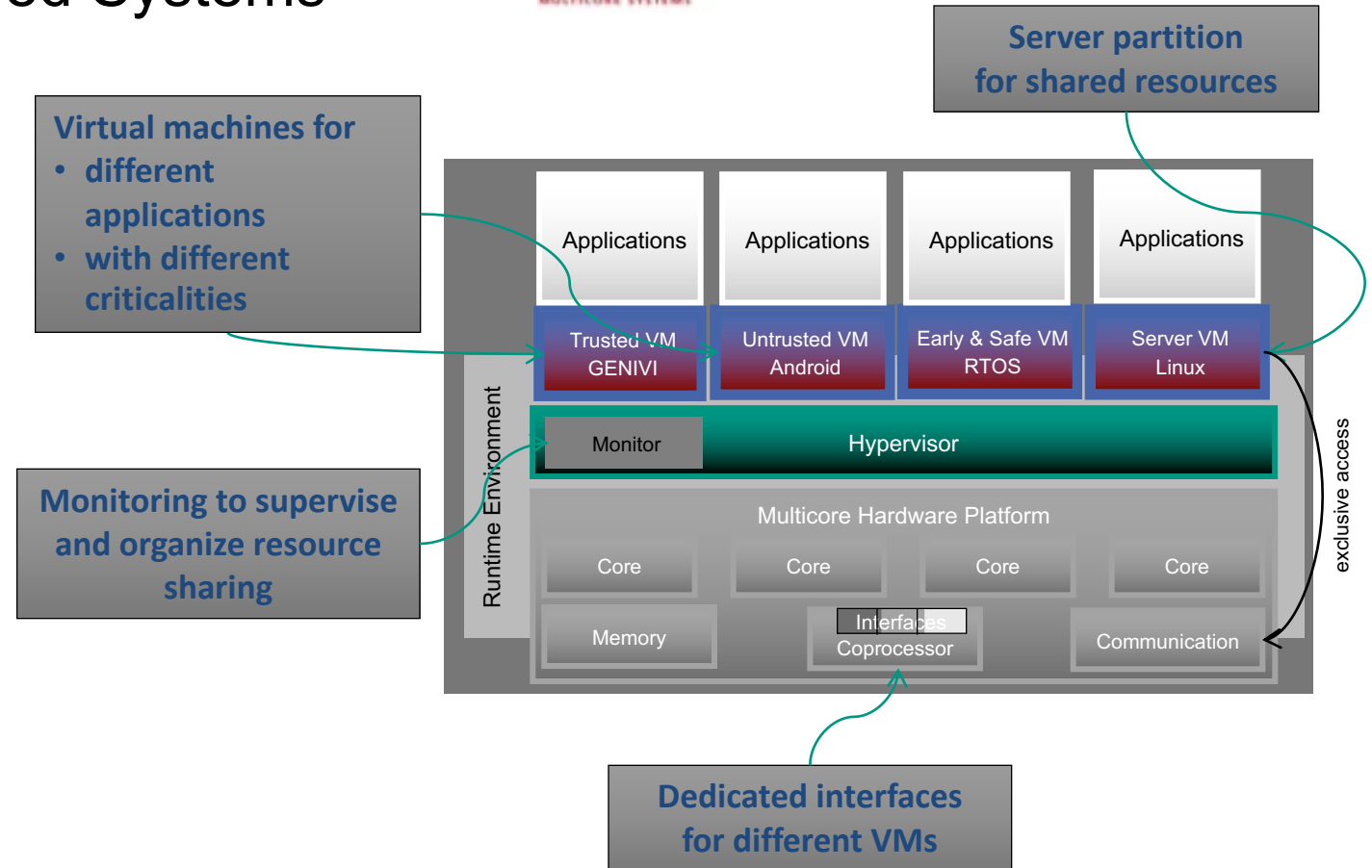
## ■ Virtualization as Key Technology for Usage of Multicore Platforms in Embedded Systems



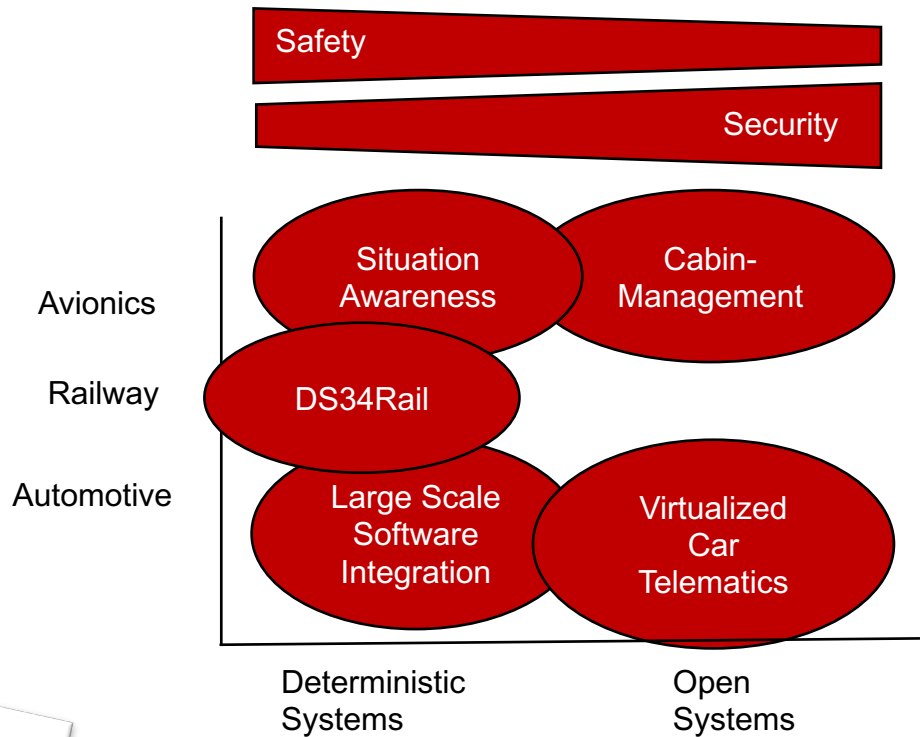
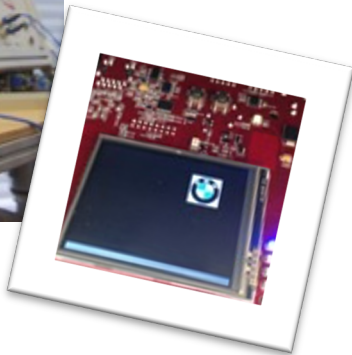
- **Consolidation** of Functions
- **Segregation** of Applications with **different criticalities** regarding Functional safety
- **Integration / re-use** of existing software

## ■ Basic principles

- Sharing and Segregation of **Computation time**
- Sharing and Segregation of **Memory**
- Secure **Communication** between Partitions



# ARAMiS Demonstrators: Automotive, Avionics, Railway



# From *ARAMiS* to *ARAMiS II*

**ARAMiS**  
proved the applicability of multicores in  
safety critical applications in principle



GEFÖRDERT VOM



**ARAMiS II**  
targets the efficient use of multicores in safety critical applications  
in practice by preparation of:



**STRUCTURED MULTICORE  
DEVELOPMENT**



**MULTICORE METHODS  
AND TOOLS**



**INDUSTRIAL PLATFORMS  
FOR MULTICORE SYSTEMS**





# Invas/C Heterogeneous Architecture

## ■ Tiled many-core architecture

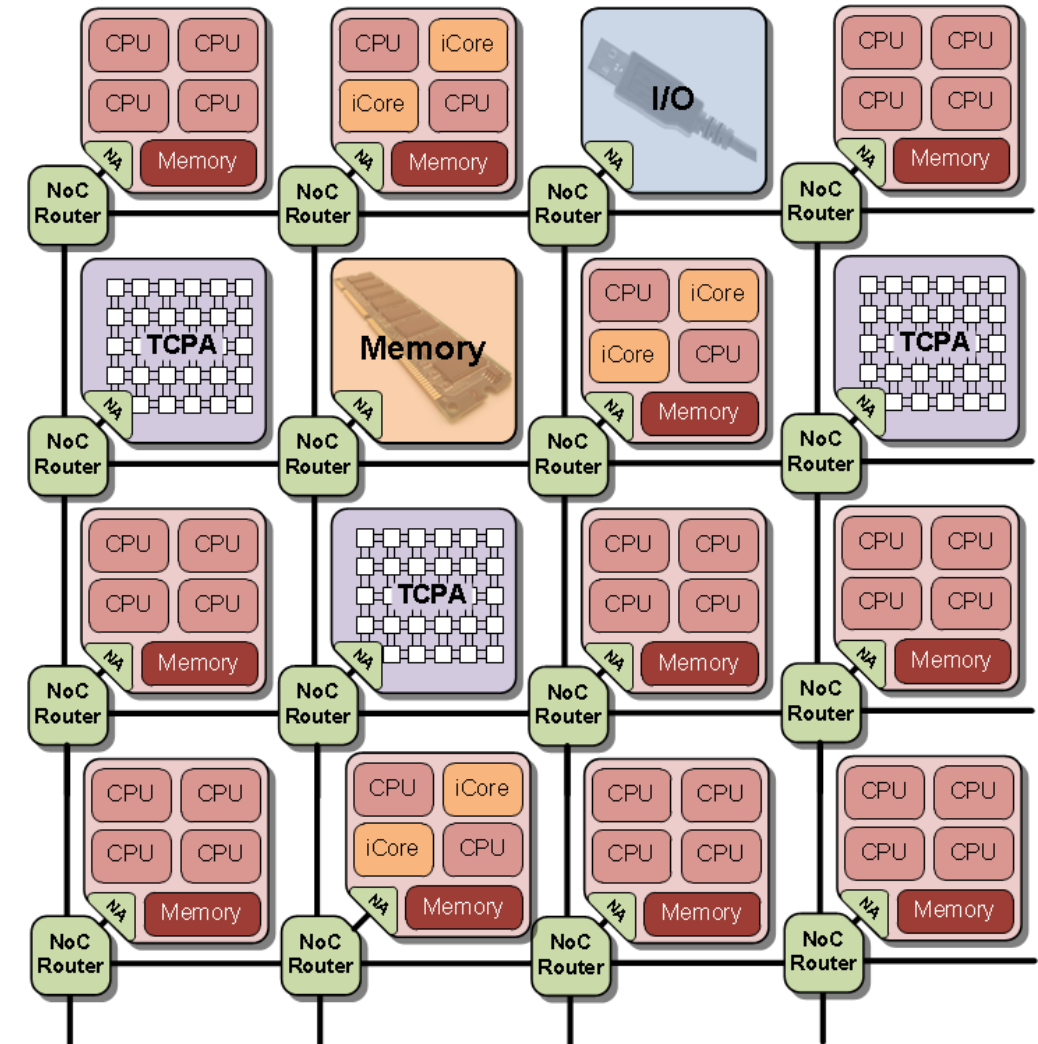
- One or more CPUs per tile
- Network-on-Chip (NoC)
- Tile Local Memory (TLM)

## ■ PGAS memory architecture

## ■ LEON3 cores and GRLIB from Cobham Gaisler

## ■ *Invasive NoC* router and network adapter offering guaranteed service connections

## ■ FPGA Prototype



# NMC and Adaptive Accelerators in *InvasIC*

## ■ Invasive Networks-on-Chip (*iNoC*)

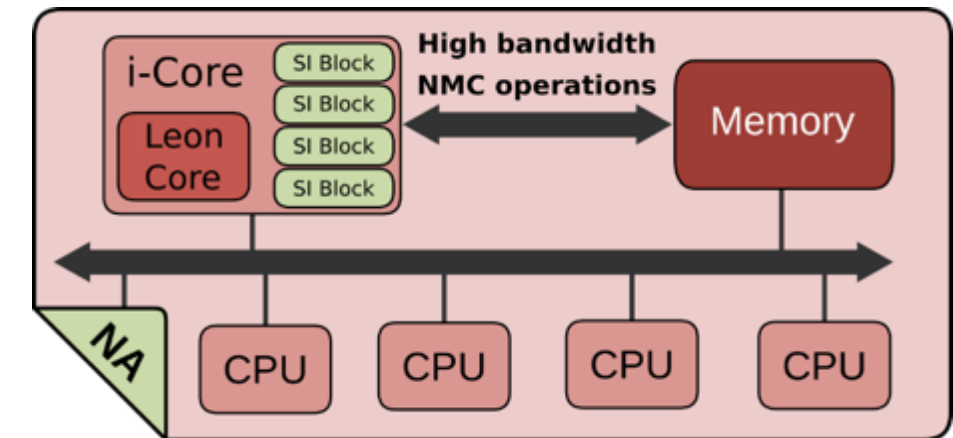
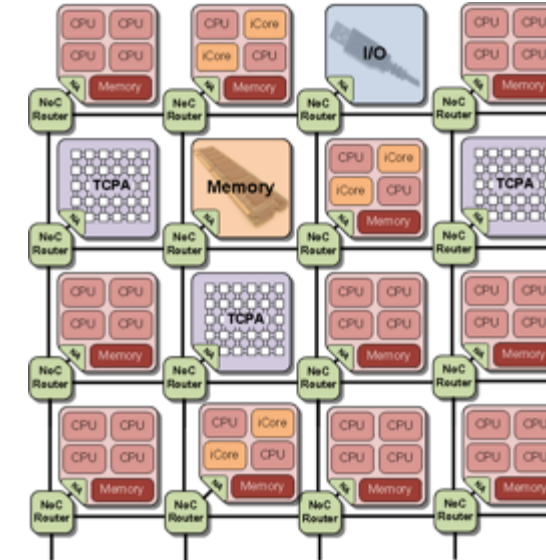
- Complex memory hierarchy
- More layers of data locality (Global Memory/Tile-Local Memory)

## ■ *i*-Core: Runtime-adaptive processor

- Integrated reconfigurable FPGA fabric
- Close to TLM with high bandwidth
- Near-memory computations

## ■ Tightly-Coupled Processor Array (*TCPA*)

- parameterizable programmable VLIW array
- accelerating computationally intensive loops

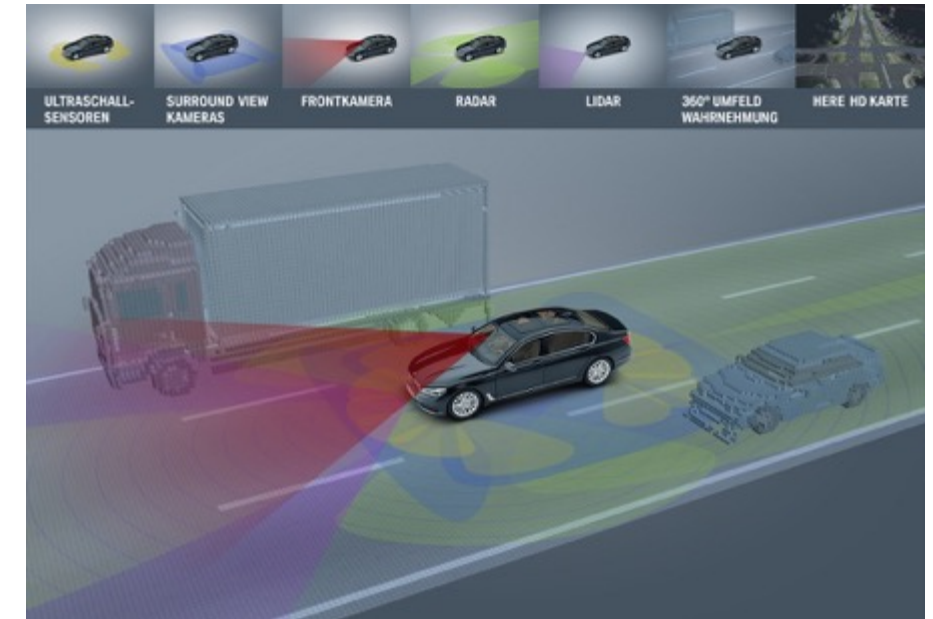




# National Project: *ZuSE-KI-mobil*



- **Goal:** Efficient **CNN accelerator SoC** for performance demanding applications such as **sensor fusion tasks** in autonomous cars
- New AI-Processor types need new approaches to guarantee **functional safety**
- Idea of a **scaleable platform** to target various fields of embedded applications



[1]

[1] <https://www.autonomes-fahren.de/bmws-bericht-zum-automatisierten-fahren/>



PARTNERS

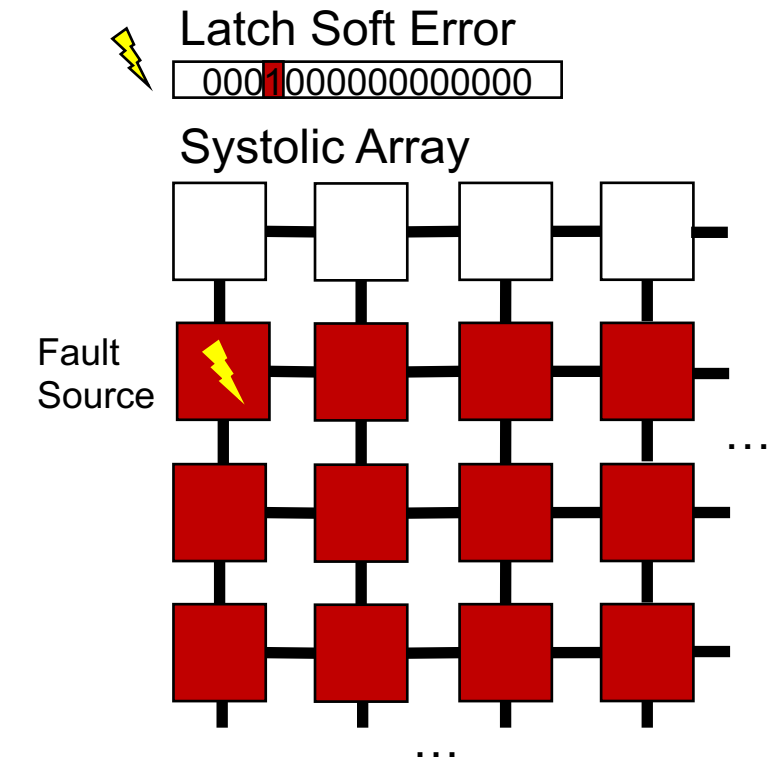


T3-technologies



# ZuSE-KI-mobil – Reliable AI Acceleration

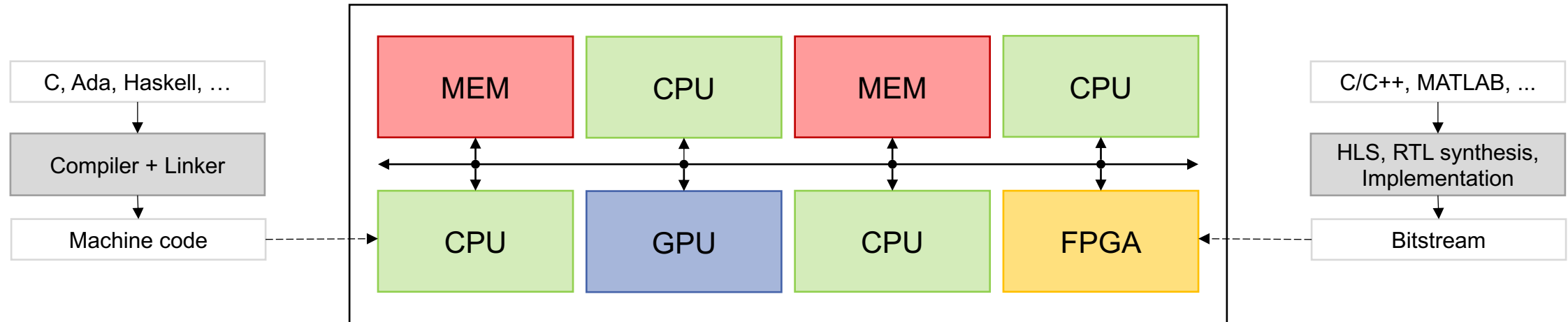
- **Transient Errors and Permanent Errors** are an increasing problem since the performance/energy demands impose the shrinking of the technology node
- **AI Accelerator** architectures tend to be vulnerable for significant **error propagation** (see picture →)
- **Research Challenges:**
  - Examine the influence of random hardware faults on state-of-the-art AI algorithms
  - **Logic BIST:** Detection of permanent hardware faults in highly parallelized accelerators
  - **Co-design, simulation and design space exploration:** Find the best performing, yet reliable hardware architecture option





# Configuration of *On-Chip Isolation Units*

- Steadily increasing complexity of COTS available *MPSoCs*:



- Functional application requirements**  $\Rightarrow$  Synthesis of low-level implementations from high-level descriptions are *state of the art*
- However:** Manual configuration of *on-chip safety & security mechanisms*  $\rightarrow$  *Reliability* requirements

# Configuration of *On-Chip Isolation Units*

- **Concept:** Model-based Spec for *On-Chip Isolation Requirements*
  - **Verification** of their **consistency** with the desired information flow policies
  - **Automatic generation** of configuration code for *on-chip isolation units*
- **Toolchain** developed as part of the *DEFEnD* project (2018 – 2021):

GEFÖRDERT VOM



Bundesministerium  
für Bildung  
und Forschung

```
// model.seg
platform plat0 {
  unit speed_sensor
  link i2c { speed_sensor, ecu.i2c1 }
  container ecu {
    generator "imx8m"
    link main { a53, i2c1 }
    unit a53
    unit i2c1
  }
}

functional func0 on plat0 {
  feature f1 on ecu.a53
  require flow f1 -> f2
  // ...
}

flow flow0 on func0 {
  // ...
}
```



Toolchain

```
// rdc_config.c
#include "rdc_config.h"

const unsigned int RDC_MDA_VALUES[] = {
  0x00000001, // Quad A53 domain assignment
  0x00000002, // M4 domain assignment
  // ...
};

const unsigned int RDC_PDAP_VALUES[] = {
  0x00000000, // GPIO1 access permissions
  0x00000010, // GPIO2 access permissions
  // ...
};
```



Deployment



**Enforcement of desired information flow  
policies (“correct by construction”)**

T. Dörr, T. Sandmann, and J. Becker. *A Formal Model for the Automatic Configuration of Access Protection Units in MPSoC-Based Embedded Systems*. Euromicro DSD 2020.

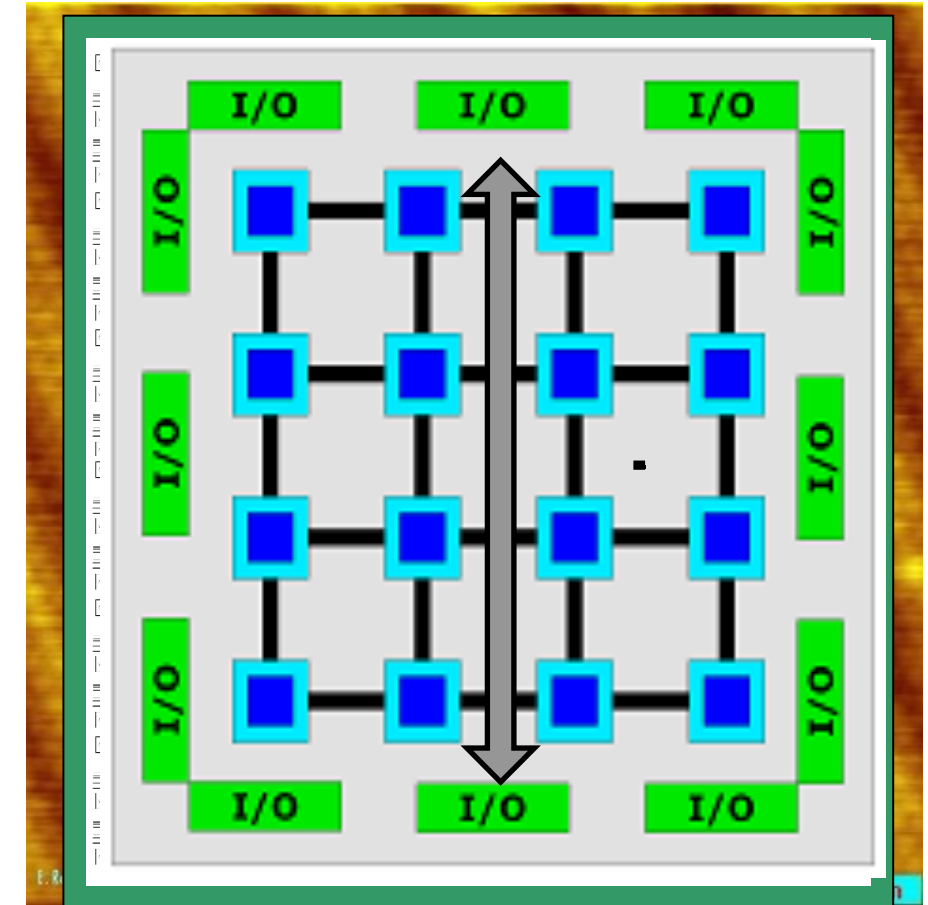
# Enabler: Embedded HPC & Reliability & Adaptivity

## Technologies & Functional Intelligence:

- Emerging & Reconfigurable Technologies:
  - **Silicon**, Optics, Nano, ... Quantum ...
- **Parallel Programming** & Computing / Data Integration
- Dynamic Networks & **Offline/Online Trade-offs**
- Correctness & **Reliability** -> **MPSoC Systems** + Dynamics
  - **Verifiable AI-Models** & Metrics
- **Safety & Security** Codesign -> **MC** + **Accelerators**
- Cross-layer System Integration -> **Adaptivity** + **Reliability**

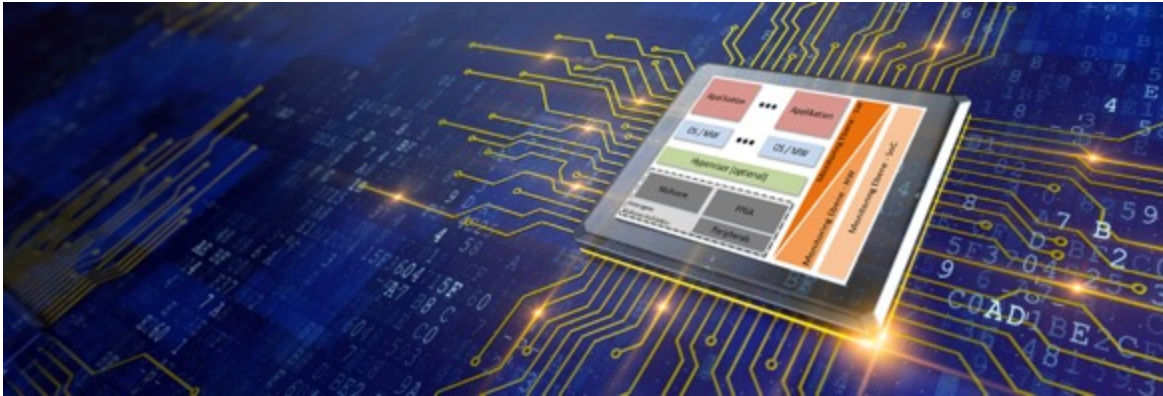
## Research, Education & Innovation

- **Hardware/Software/Architecture/Algorithms Codesign** & **AI**
- Tools & Methods for **Programming** & **Resource Management**
- Challenges: **Embedded Integration of AI, HPC & Reliability** -> **Open Hardware & Decentralization ...**



# Thanks for your attention!

... Silicon will live for ever!



## Contact:

**Prof. Dr.-Ing. Dr.h.c. Jürgen Becker**

Karlsruher Institut für Technologie (KIT)

Institut für Technik der  
Informationsverarbeitung (ITIV)

Engesserstr. 5

76131 Karlsruhe

[www.itiv.kit.edu](http://www.itiv.kit.edu)

Email: [becker@kit.edu](mailto:becker@kit.edu)

