Models of Architecture

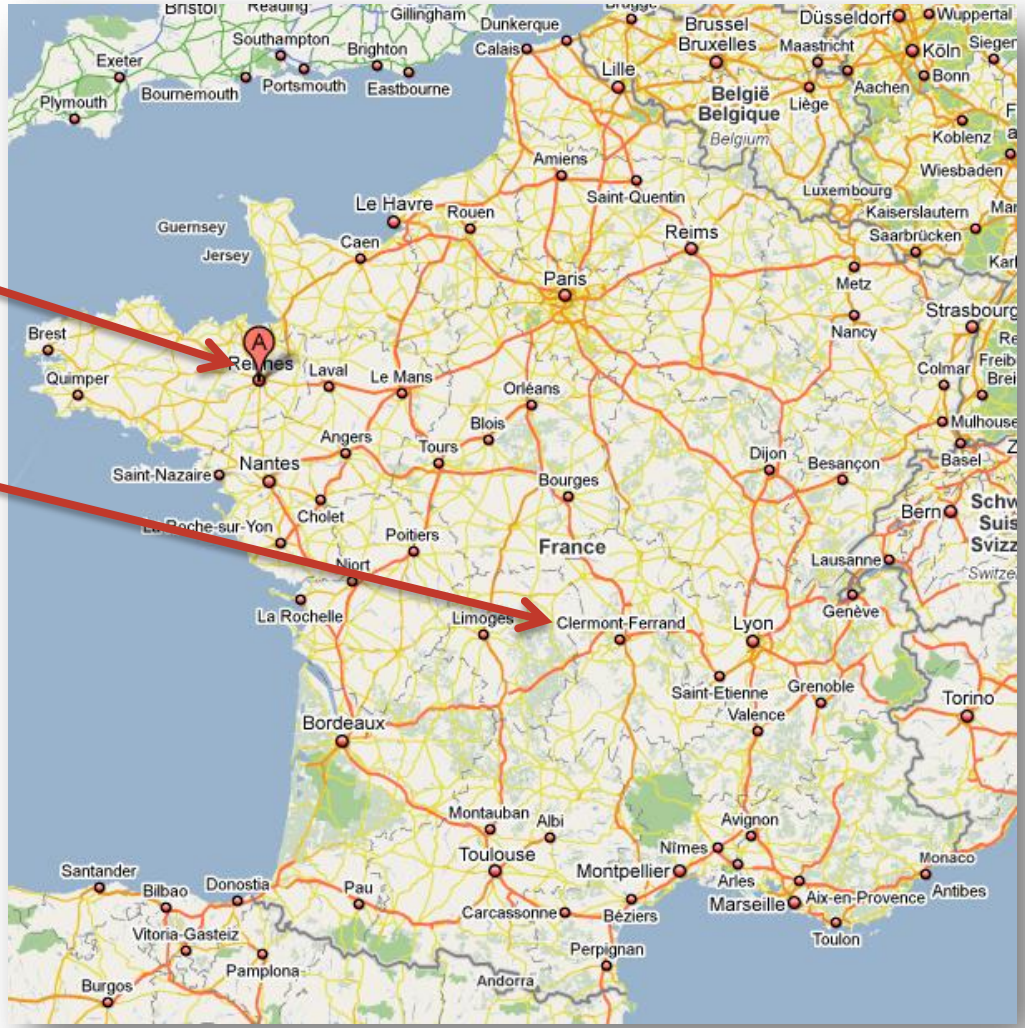
Maxime Pelcat
INSA Rennes, IETR, Institut Pascal

Nokia Bell Labs 2018

This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 732105: CERBERO.

INSA Rennes – IETR VAADER

- INSA Rennes
- IETR VAADER
- Institut Pascal



Models of Architecture

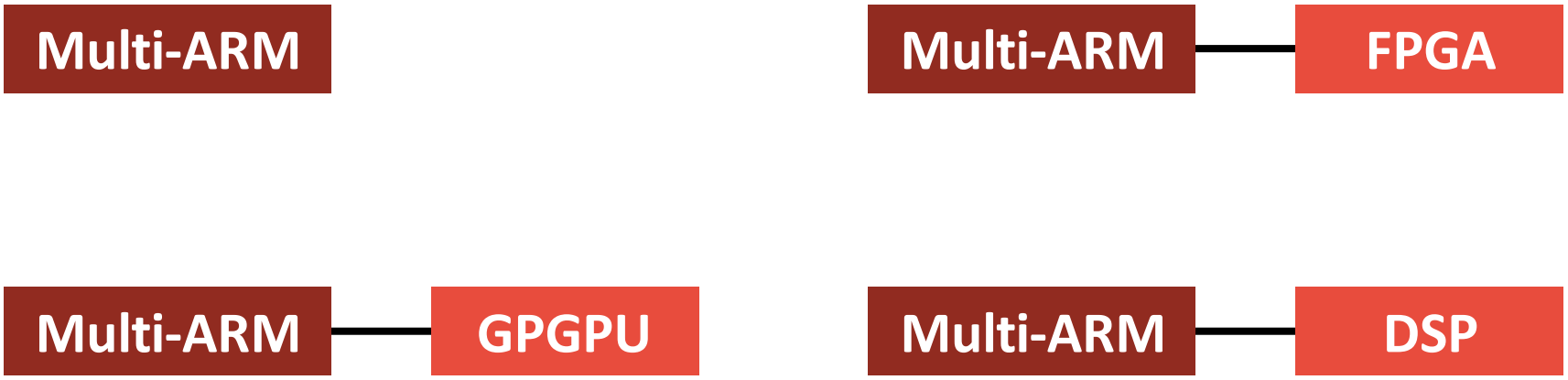
- Abstracting computational architecture to
  - **Predict performance**
  - **Support current hardware evolutions**

Motivation: architecture evolution

- Hardware Architectures are becoming
  - **More complex**
  - **More heterogeneous**
  - **More High Performance embedded Computing (HPeC)**
    - Embedded deep learning, near-sensor computing, fog computing, edge computing, many-cores, etc.
    - Real-time constraints, stream processing applications

Motivation: HPeC architectures

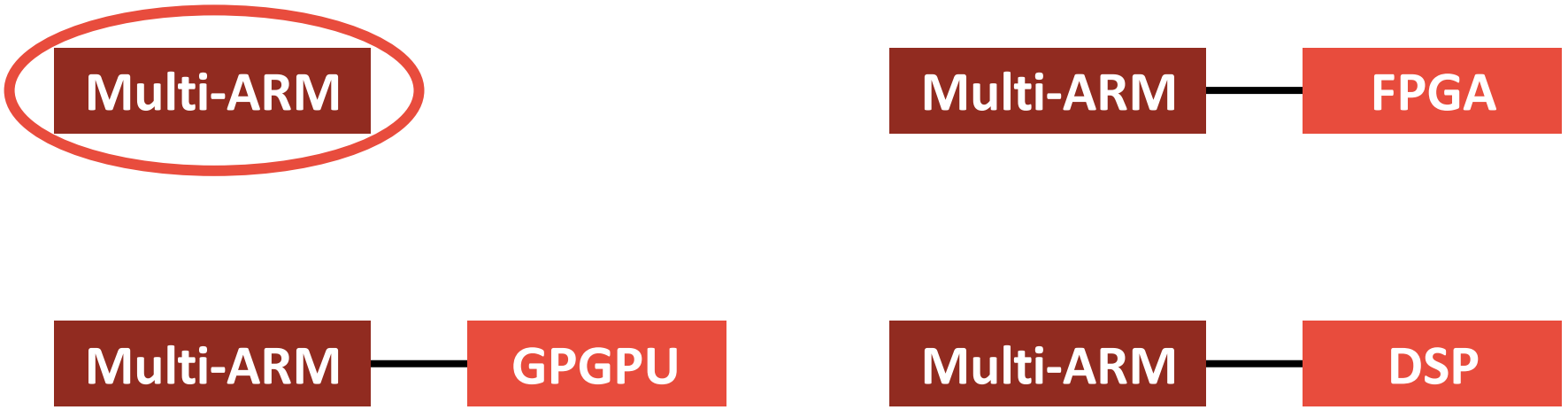
- Let's look at ARM-based HPeC
  - **Let us consider 4 heterogeneous solutions**
    - ARM = control path + some of the data path
    - in red: data path



INSA IETR
Horizon 2020
European Union Funding
for Research & Innovation
CERBERO

Motivation: HPeC architectures

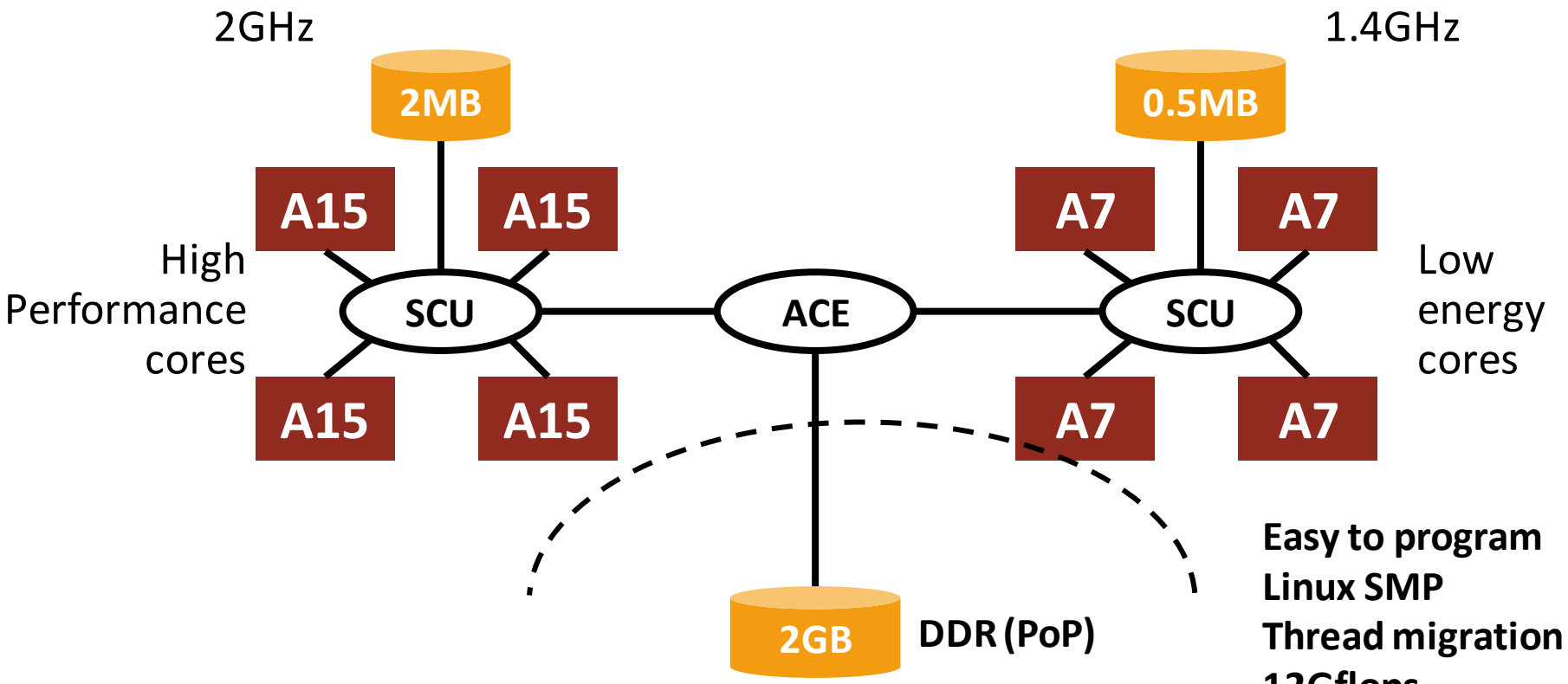
- Let's look at ARM-based HPeC



INSA IETR Horizon 2020 European Union Funding for Research & Innovation CERBERO

Motivation: HPeC architectures

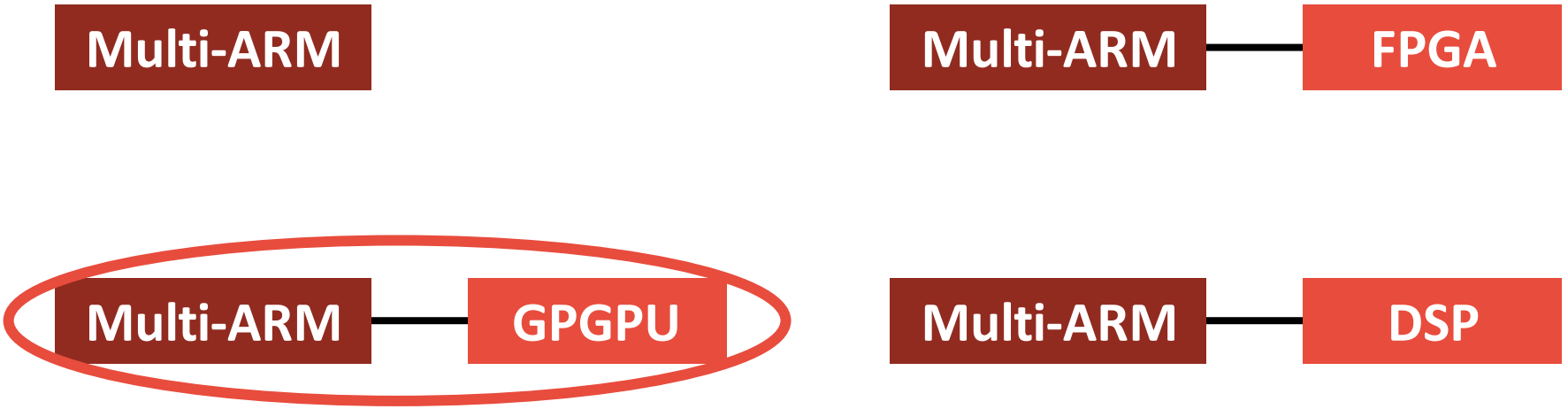
- ARM big.LITTLE: Samsung Exynos 5422



INSA IETR Horizon 2020 European Union Funding for Research & Innovation CERBERO

Motivation: HPeC architectures

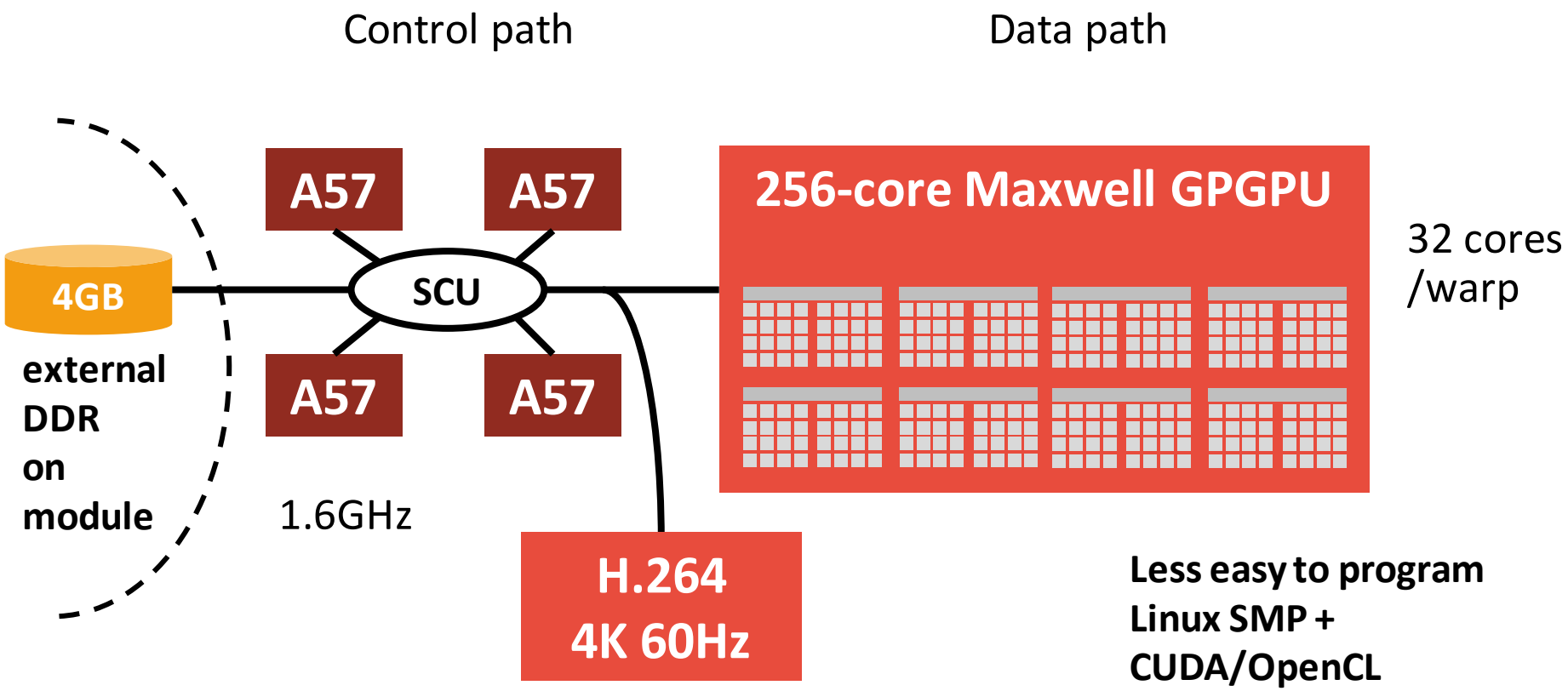
- Let's look at ARM-based HPeC



INSA IETR Horizon 2020 European Union Funding for Research & Innovation CERBERO

Motivation: HPeC architectures

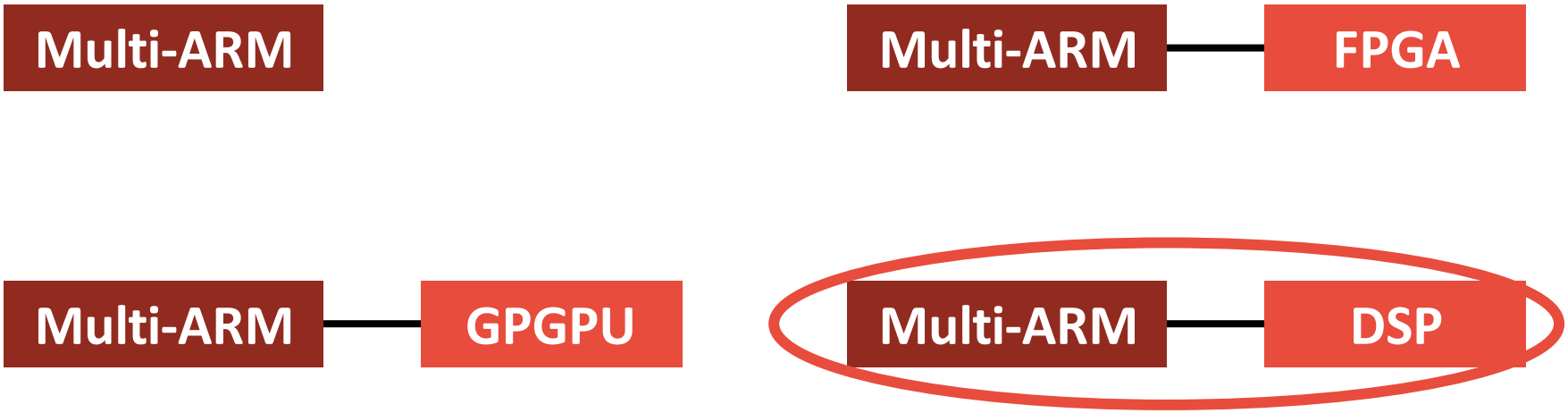
- Multi-ARM + GPGPU: Nvidia Jetson TX1 module



INSA IETR Horizon 2020 European Union Funding for Research & Innovation CERBERO

Motivation: HPeC architectures

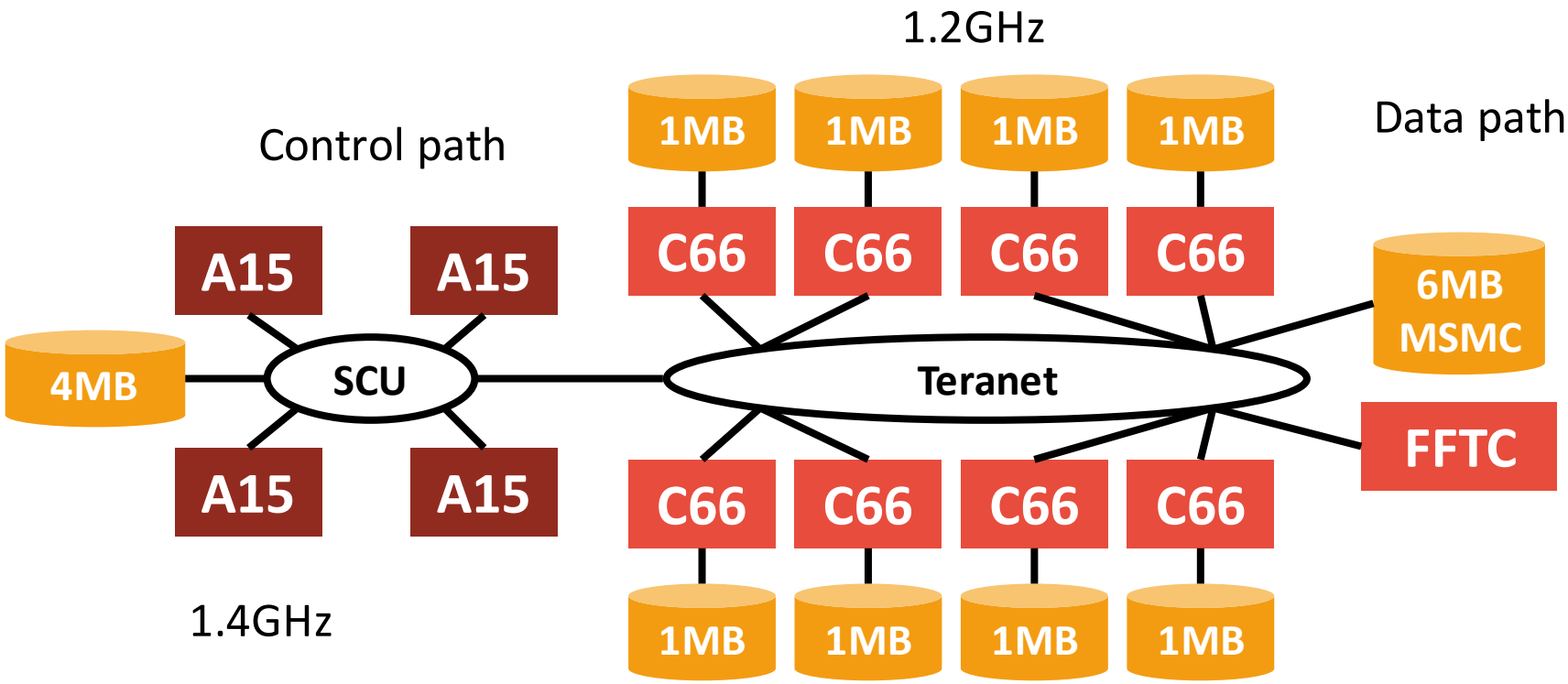
- Let's look at ARM-based HPeC



INSA IETR Horizon 2020 European Union Funding for Research & Innovation CERBERO

Motivation: HPeC architectures

- Multi-ARM + DSP: Texas Instruments Keystone II TCI6638K2K

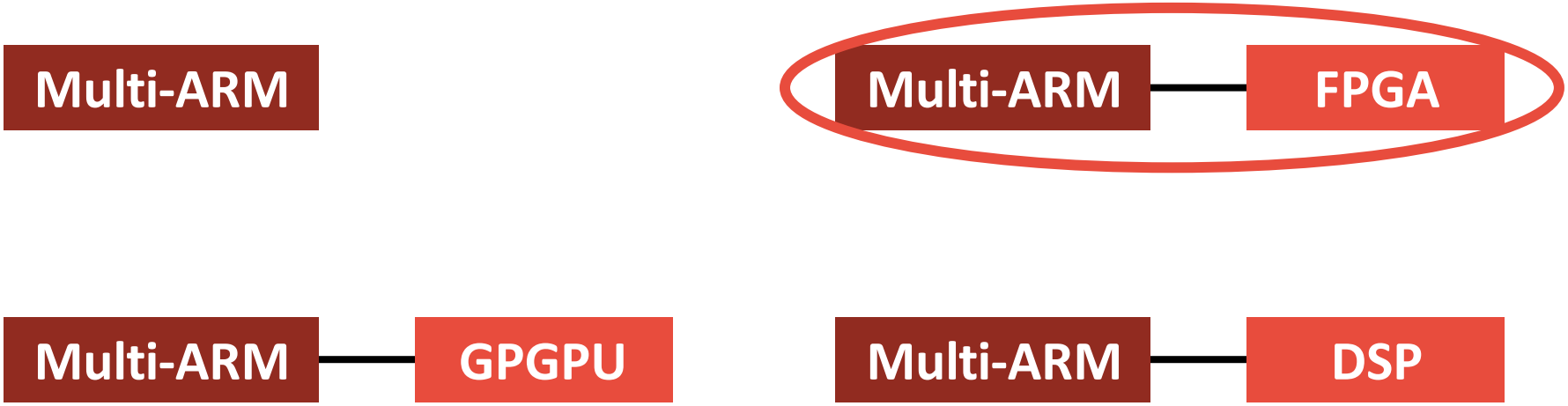


Difficult to program (well)
Linux SMP +
Open Event Machine
160 Gflops
<15W

INSA IETR Horizon 2020 European Union Funding for Research & Innovation CERBERO

Motivation: HPeC architectures

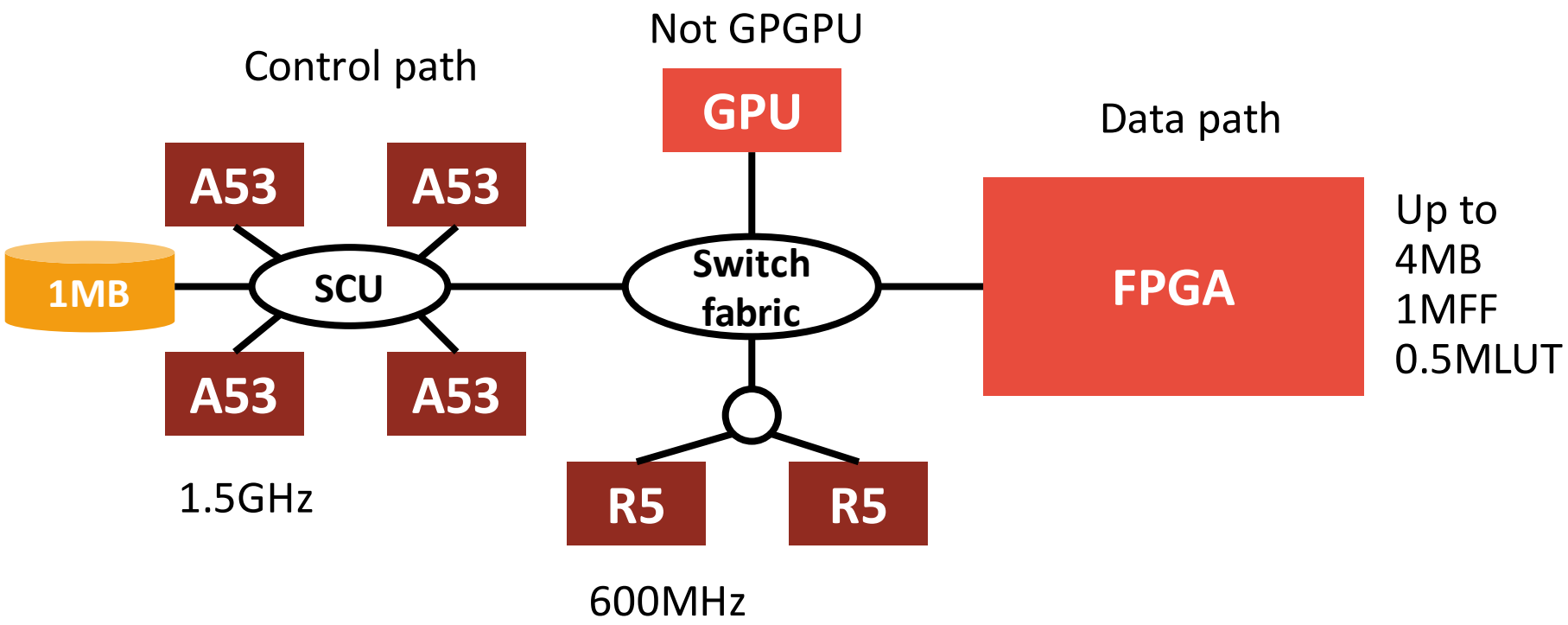
- Let's look at ARM-based HPeC



INSA IETR
Horizon 2020
European Union Funding
for Research & Innovation
CERBERO

Motivation: HPeC architectures

- Multi-ARM + FPGA: Xilinx Zynq Ultrascale +



More difficult to program (well)
Linux SMP + HLS or HDL

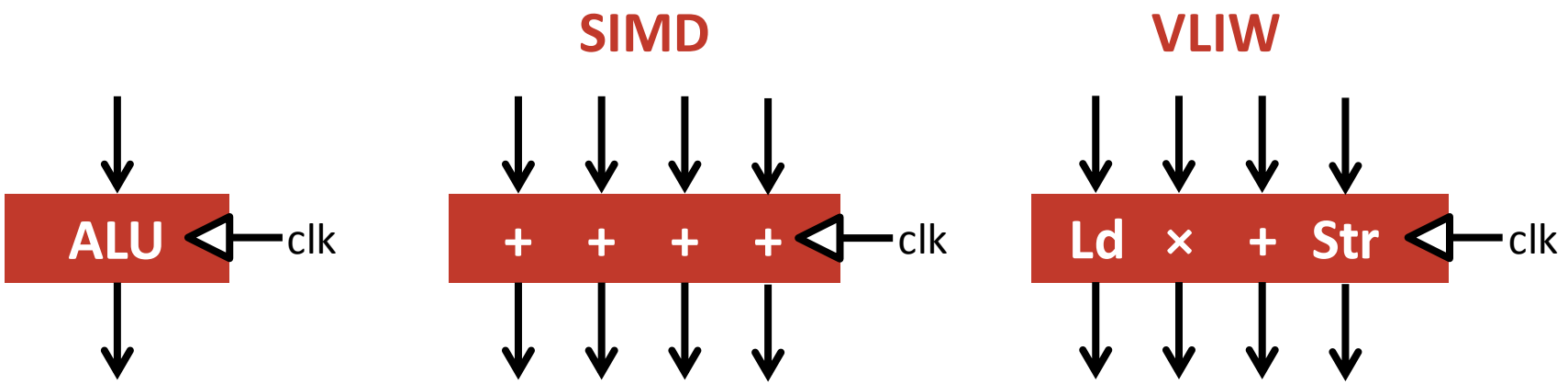
INSA IETR Horizon 2020 European Union Funding for Research & Innovation CERBERO

Motivation: HPeC architectures

- Current trends
  - **FPGAs are gaining importance: what about flops?**
  - **Adding video/image accelerators**
    - Video Compression: H.264/AVC, H.265/HEVC, etc.
    - AI: For tensor applications → reach 1Tops/W
  - **RISC-V as an open HW competitor to ARM**

Motivation: architecture evolution

- Towards more complexity
  - **More cores, hierarchies of clusters**
  - **Heteronegeneity, Interconnect complexity**
- Reminds intra-core modifications in XXth

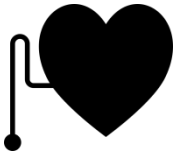


INSA IETR Horizon 2020 European Union Funding for Research & Innovation CERBERO

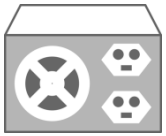
Motivation: architecture evolution

- But there are some differences between intra-core and inter-core parallelism
  - **At coarse grain, PEs communicate asynchronously**
  - **There is no (or less) centralized processing decision**
  - **There is no performance portability (nothing equivalent to C-to-VLIW compilers)**
- How can/should we manage this HW complexity?
  - **Can we predict performance at design time? How?**

System Objectives

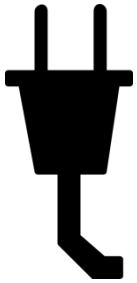


Reliability



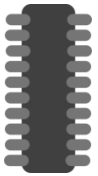
Peak Power

T°C



Energy

Performance



Memory

Security

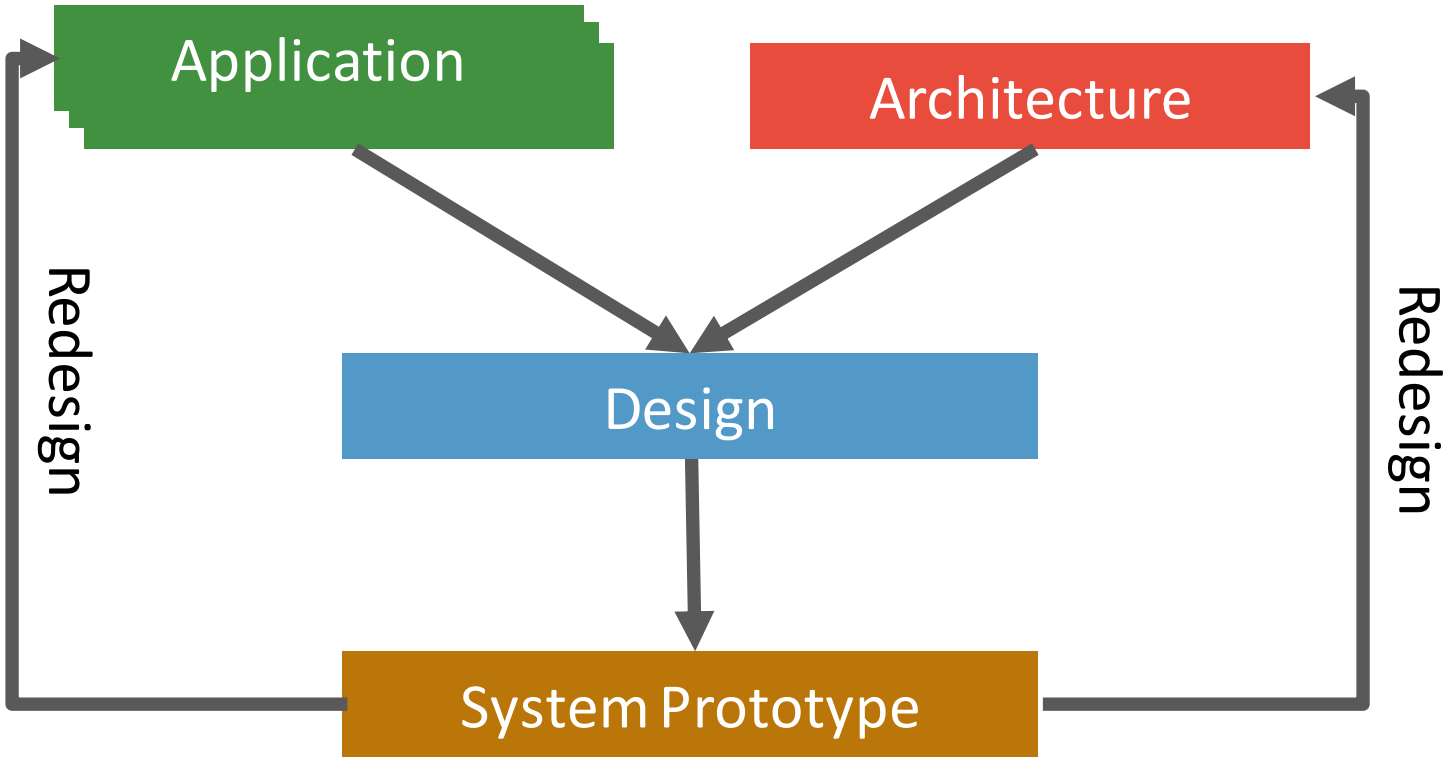
$

Unit Cost

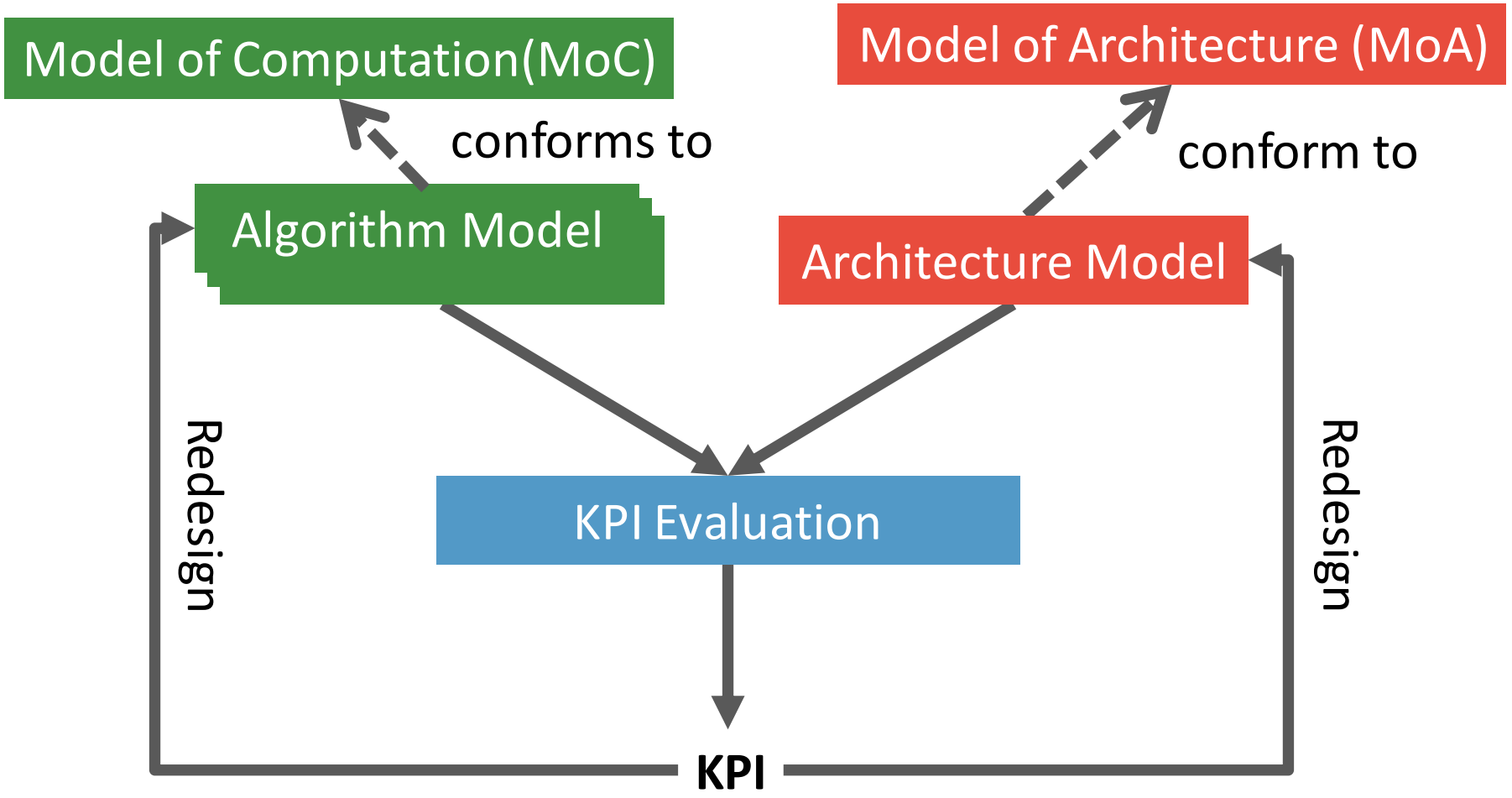
$

Maintenance Cost

System Design: Y-Chart

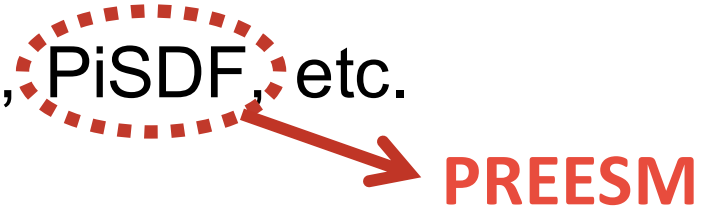


Model-Based Design



On MoC Side: Many Results

- #EdwardALee, #ProgrammingParadigms
- Discrete Event MoCs
- Finite State Machines → imperative languages
- Functional MoCs
- Petri Nets
- Dataflow MoCs SDF, CSDF, IDF, IBSDF, PSDF, SPDF, PiSDF, etc.



Dataflow MoCs Case

And they are not all here…

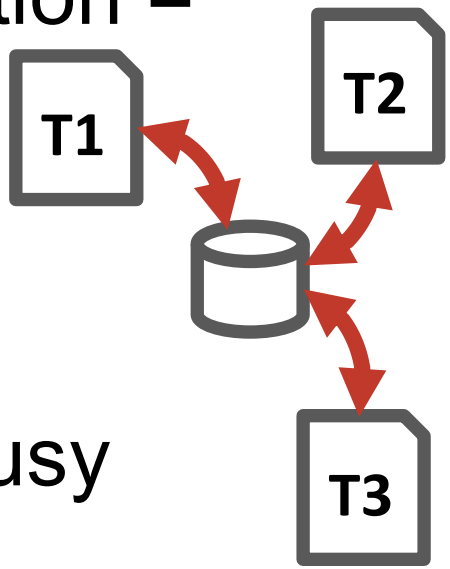
Feature	SDF	ADF	IBSDF	DSSF	PSDF	PiSDF	SADF	SPDF	DPN	KPN
Expressivity	Low				Med.		Turing complete			
Hierarchical			X	X	X	X				
Compositional			X	X		X				
Reconfigurable					X	X	X	X	X	X
Statically schedulable	X	X	X	X						
Decidable	X	X	X	X	(X)	(X)	X	(X)		
Variable rates		X			X	X	X	X	X	X
Non-determinism							X	X	X	

SDF: Synchronous Dataflow
ADF: Affine Dataflow
IBSDF: Interface-Based Dataflow
DSSF: Deterministic SDF with Shared Fifos
PSDF: Parameterized SDF
PiSDF Parameterized and Interfaced SDF
SADF: Scenario-Aware Dataflow
SPDF: Schedulable Parametric Dataflow
DPN: Dataflow Process Network
KPN: Kahn Process Network

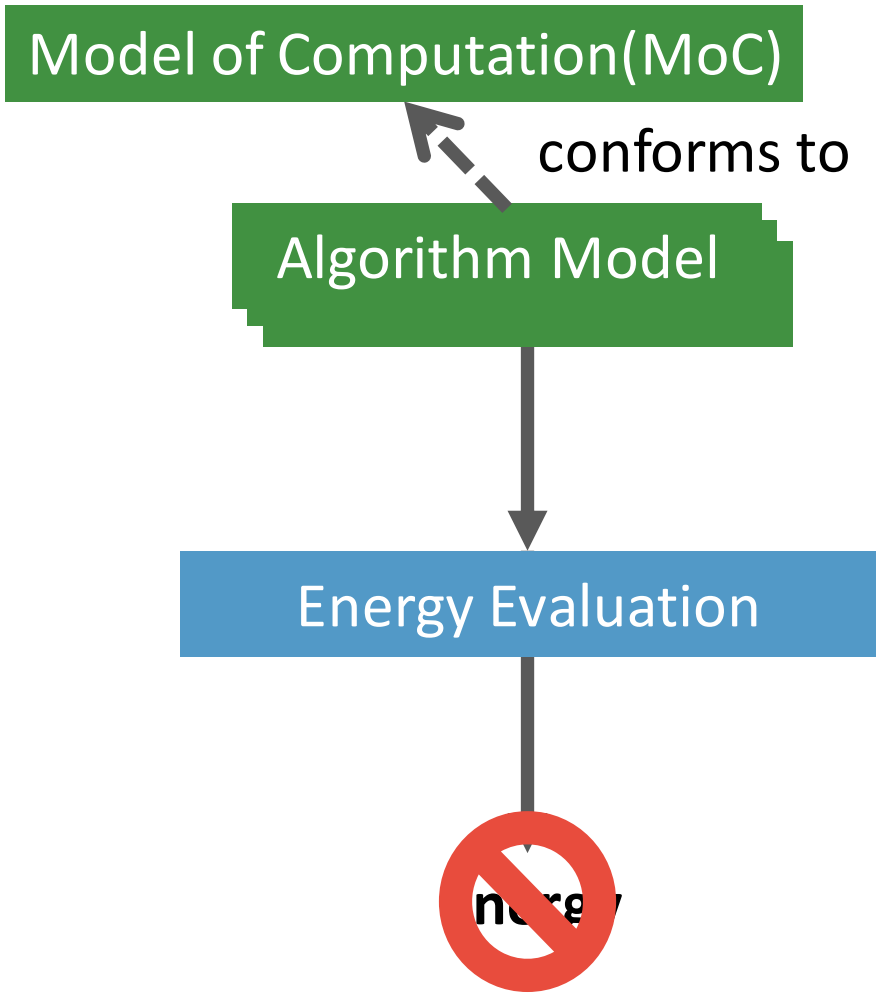
INSA IETR Horizon 2020 European Union Funding for Research & Innovation CERBERO

But Still a Lot to Do

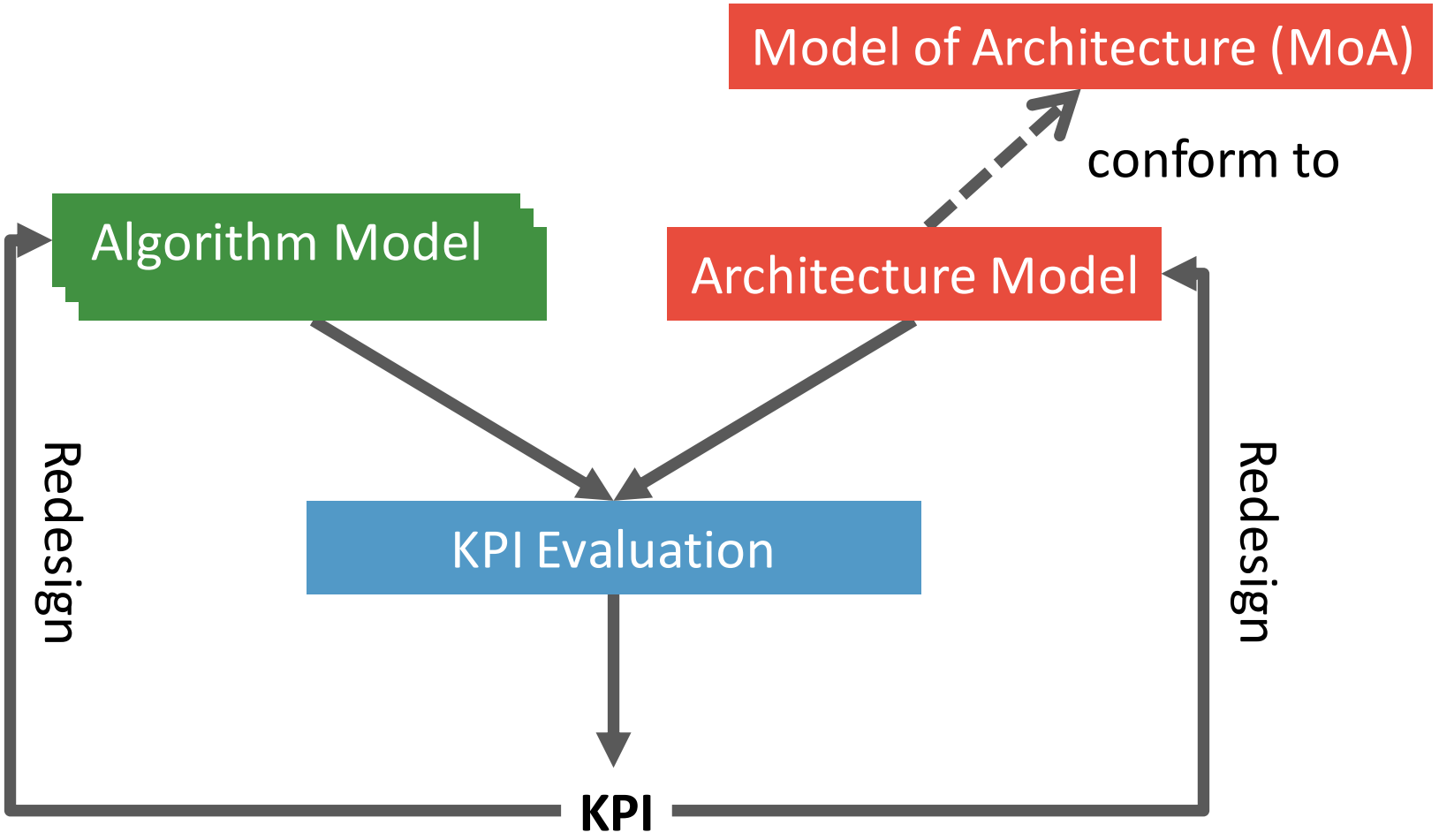
- on Real-Time Multicore systems especially
- Usually, RT application specification =
  - **Multiple tasks sharing resources**
  - **Activation periods or triggering events**
- Objective = keeping resources busy



MoCs are not sufficient



Models of Architecture



Problem: Predict System Quality

- How to predict a system « quality » ?
  - **Efficiently (simple procedure)**
  - **Early (from abstract models)**
  - **Accurately (with a good fidelity)**
  - **With reproducibility (same models = same prediction)**

Model of Architecture

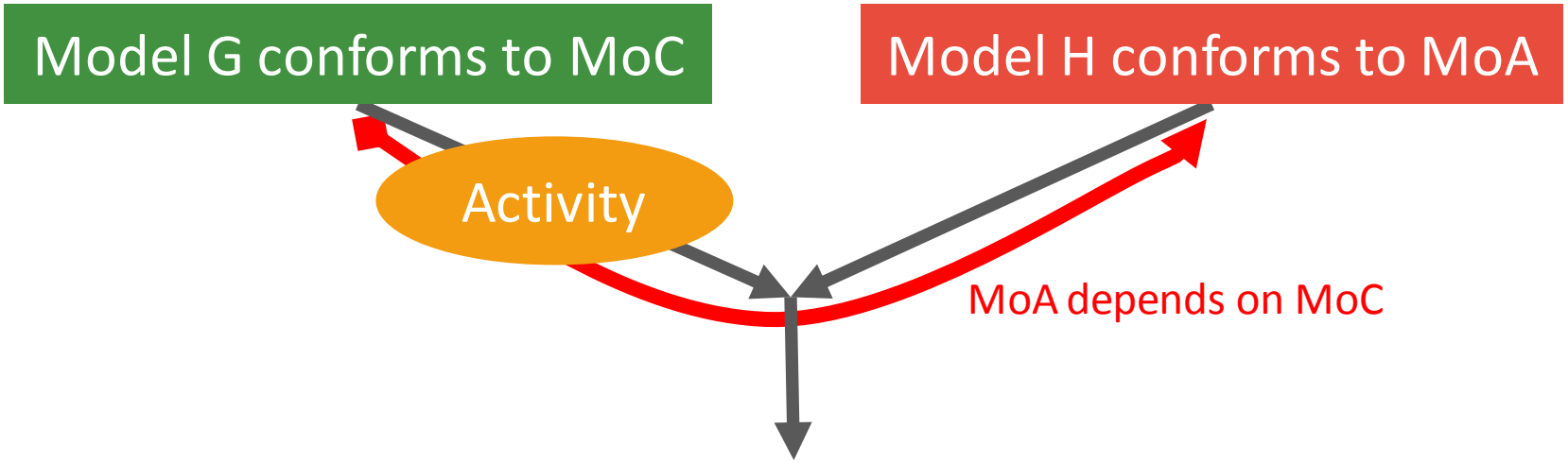
- Definition
  - **Model of a system Non-Functional Property**
  - **Application-independent**
  - **Abstract**
  - **Reproducible**

Pelcat, M; Mercat, A; Desnos, K; Maggiani, L; Liu, Y; Heulot, J; Nezan, J-F; Hamidouche, W; Ménard, D; Bhattacharyya, S (2017) "Reproducible Evaluation of System Efficiency with a Model of Architecture: From Theory to Practice", IEEE TCAD.

Model of Architecture

Model	Reproducible	Application-independent	Abstract
AADL	☑	☒	☒
MCA SHIM	☒	☒	☒
UML MARTE	☒	☑ / ☒	☒
AAA	☒	☑	☒
CHARMED	☒	☑	☒
S-LAM	☒	☑	☒
MAPS	☒	☑	☒
LSLA	☑	☑	☑

Model of Architecture

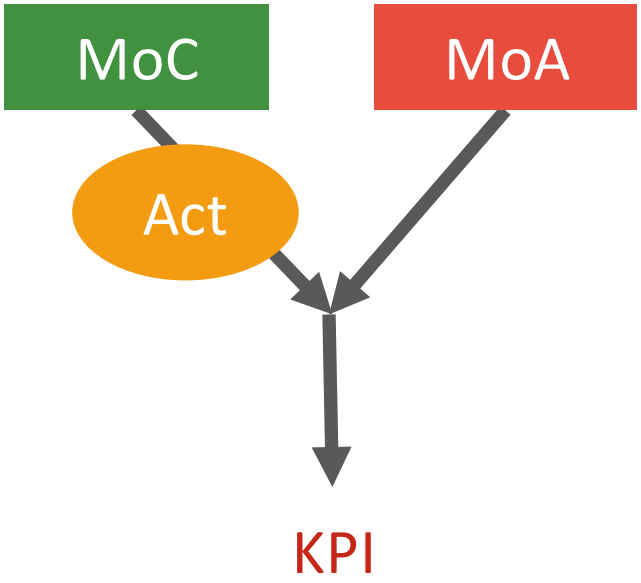


One and always the same quality evaluation

Reliability Power
Energy Performance Memory
Security T°C
Cost

= **MoA(activity(MoC(application)))**

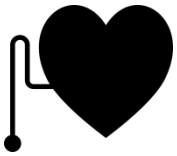
Model of Architecture



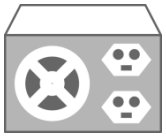
LSLA: First MoA

- LSLA = Linear System-Level Architecture Model
- Motivated by the additive nature of energy consumption

System Objectives



Reliability



Peak Power

T°C

Energy

Performance

Memory

Security

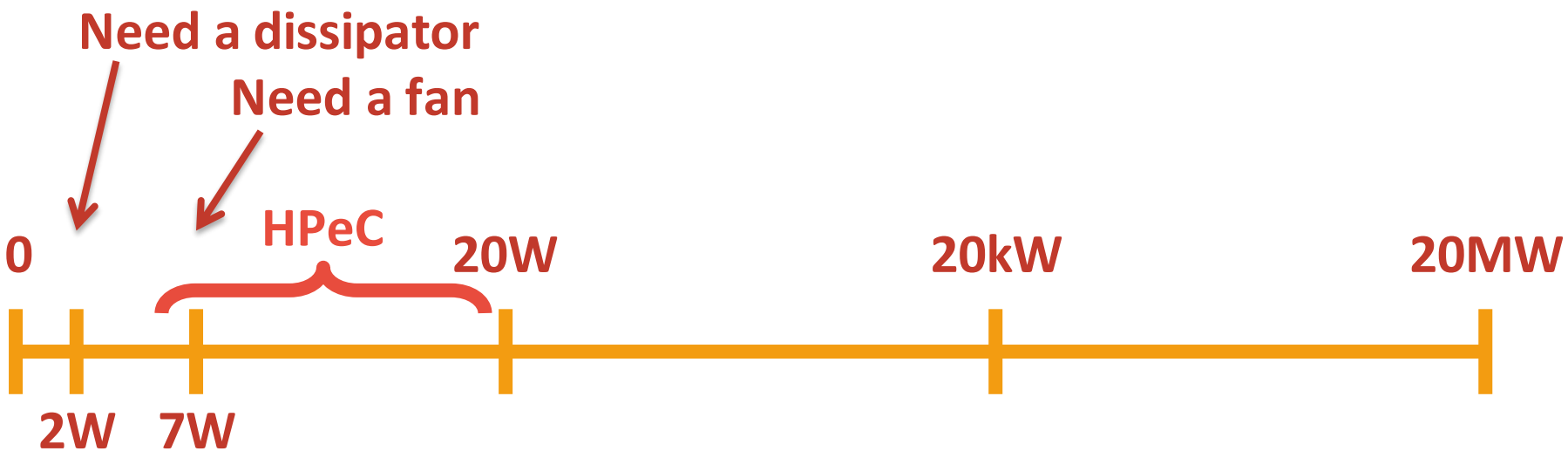
$

Unit Cost

$

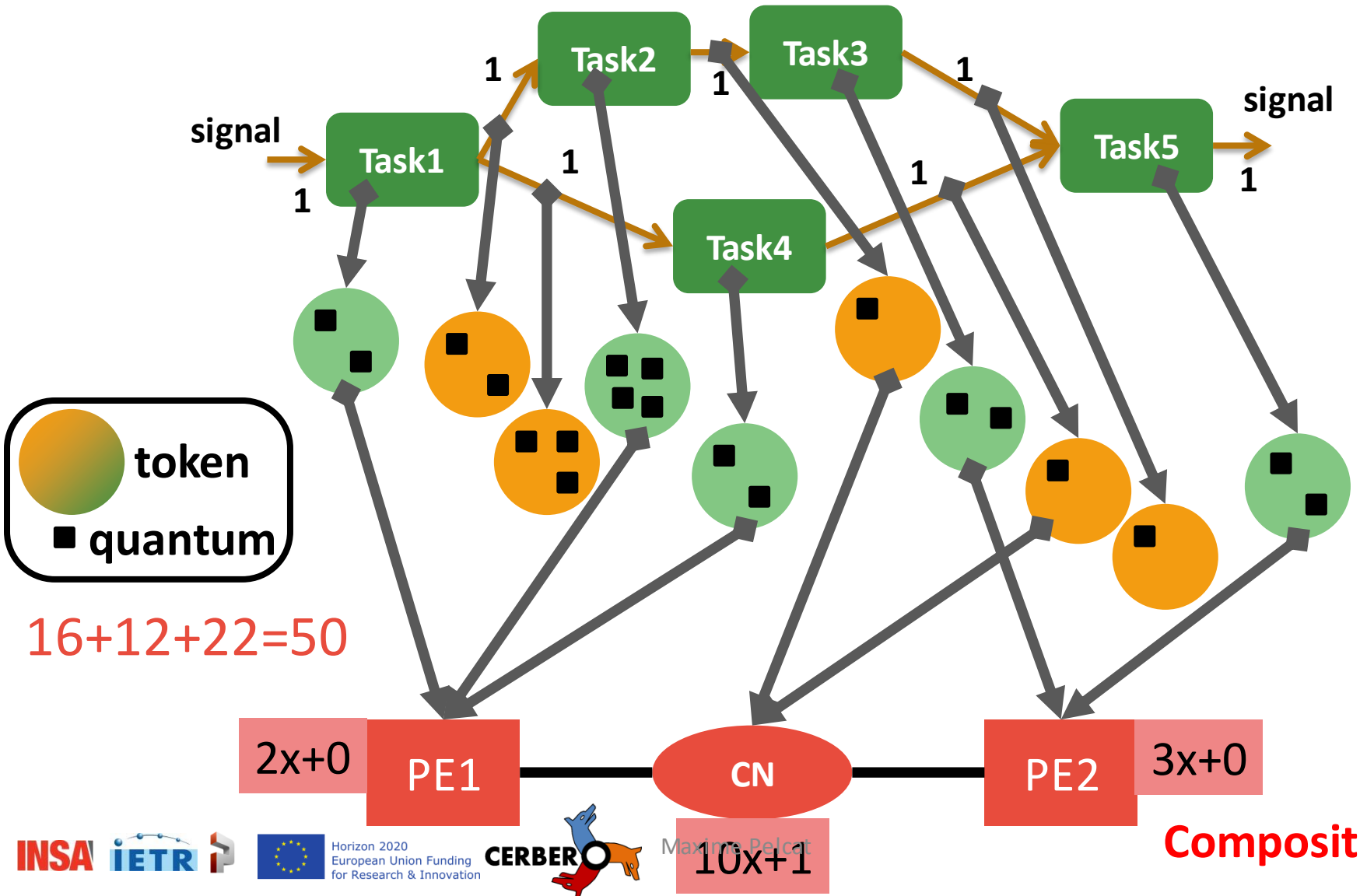
Maintenance Cost

Energy/Power Define Architecture



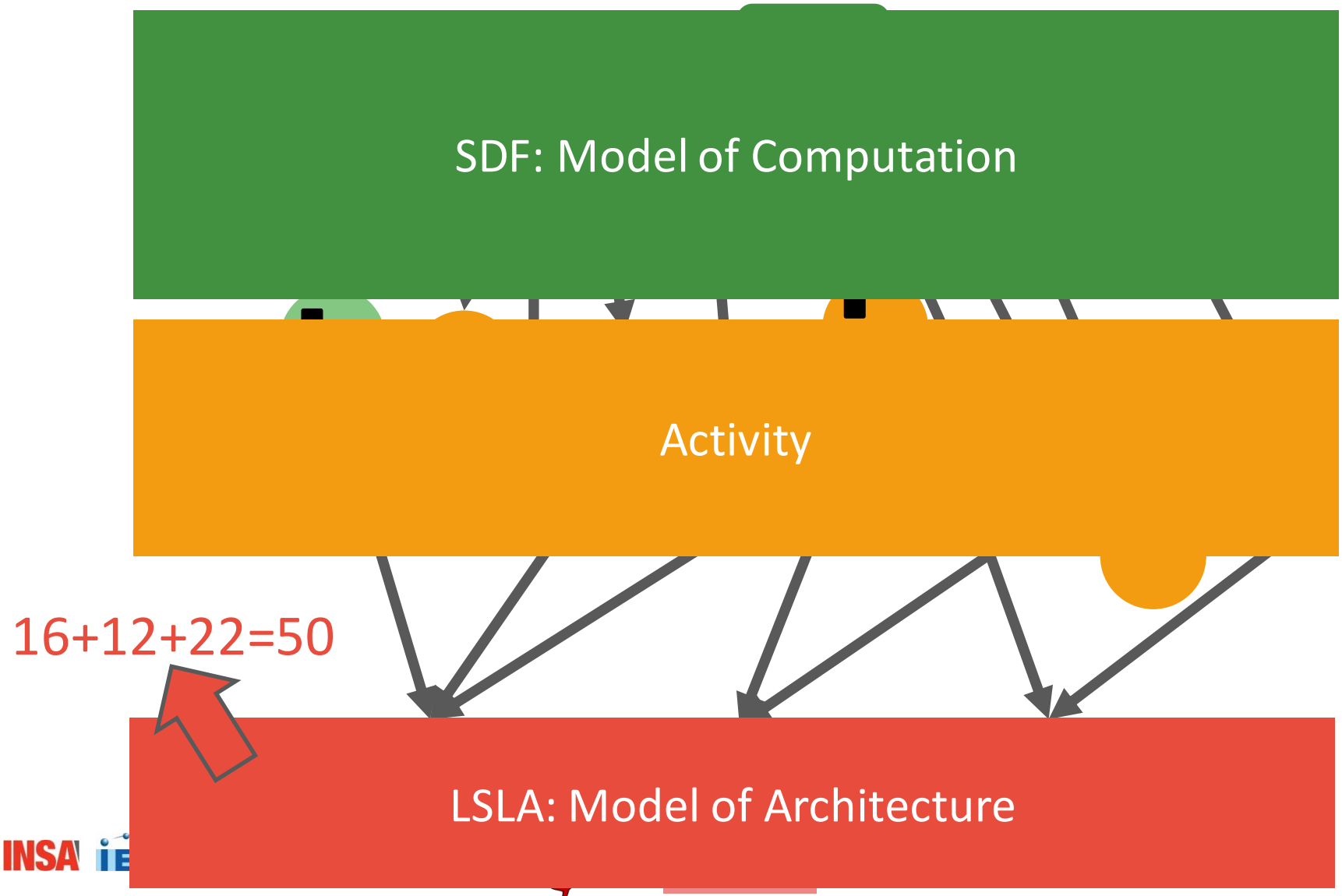
INSA IETR
Horizon 2020
European Union Funding
for Research & Innovation
CERBERO

LSLA Model of Architecture



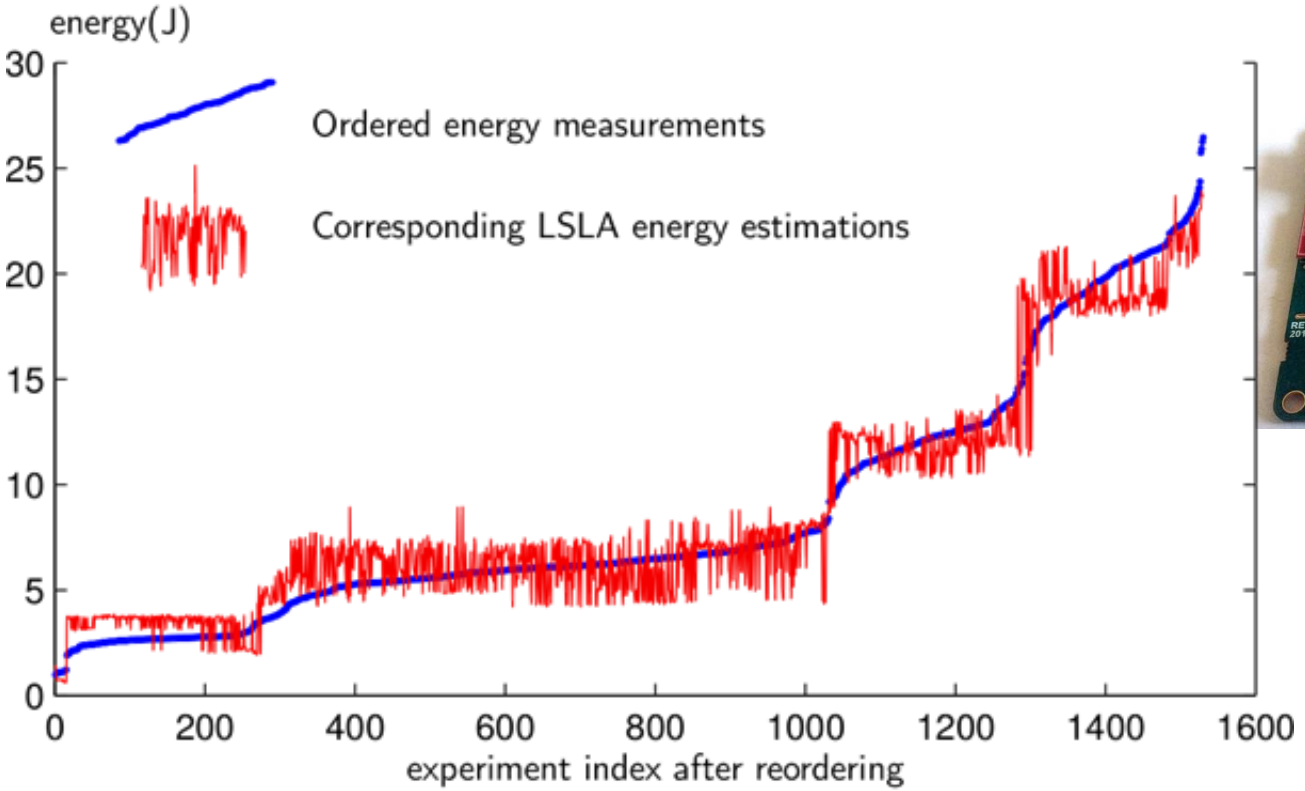
Compositional

LSLA Model of Architecture



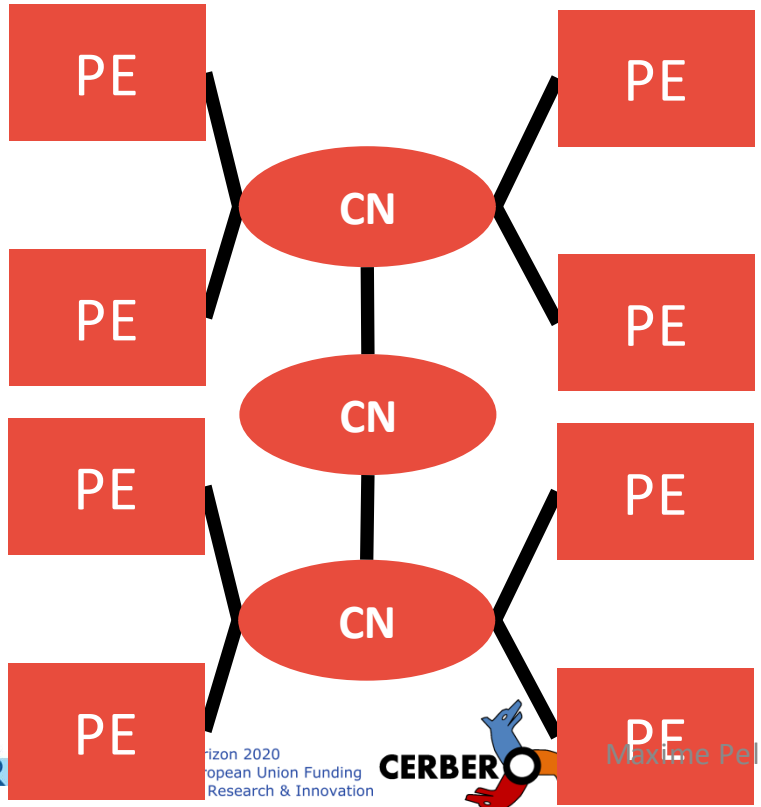
LSLA MoA for Energy Prediction

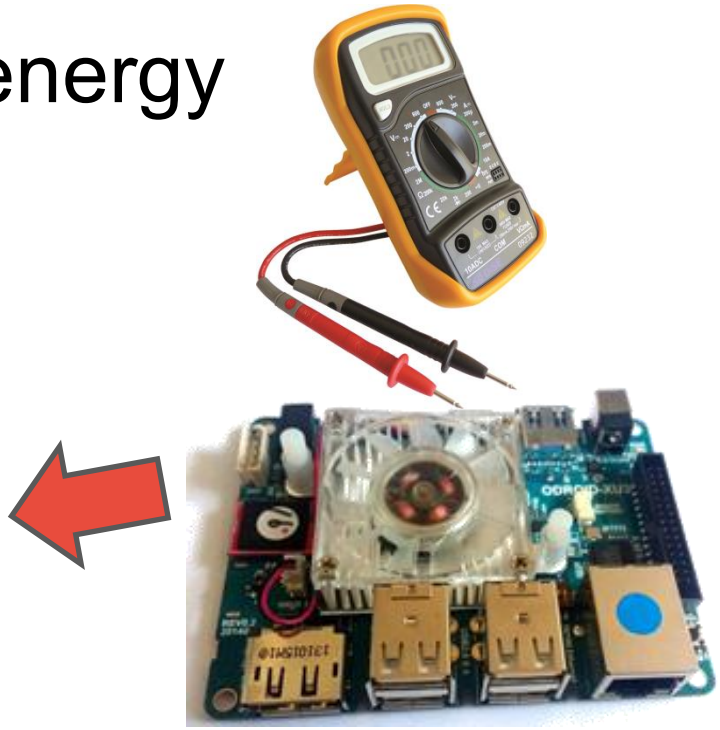
- 86% of fidelity on octo-core ARM ☺



LSLA MoA for Energy Prediction

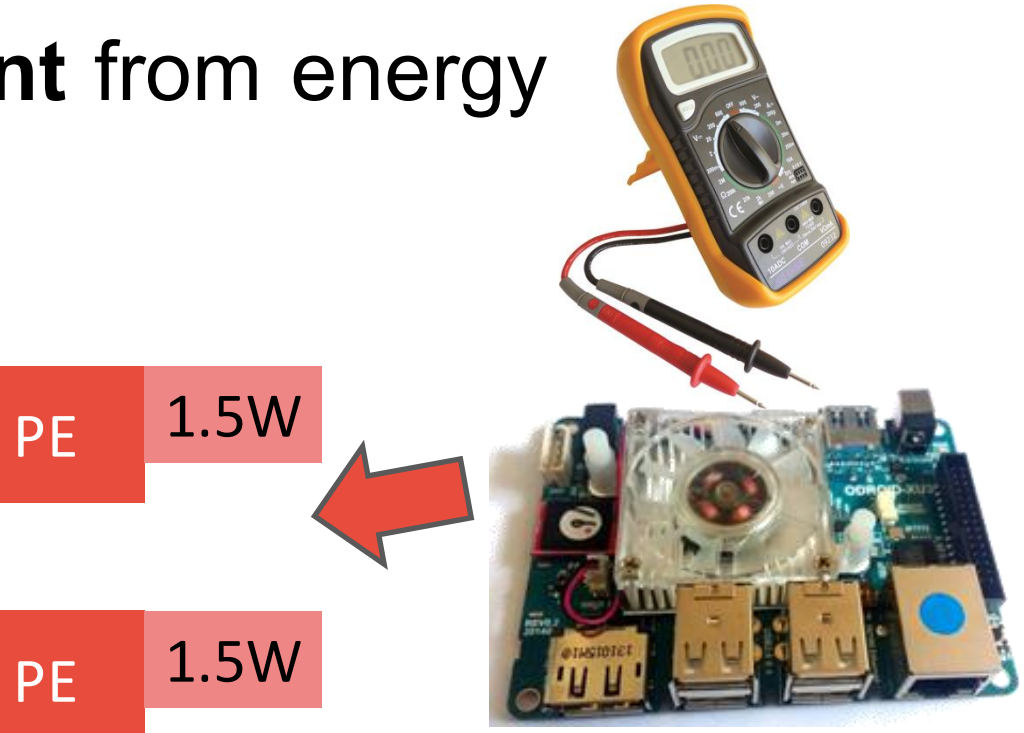
- The model is learnt from energy measurements





LSLA MoA for Energy Prediction

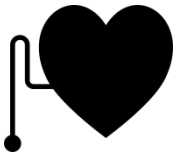
- The model is **learnt** from energy measurements



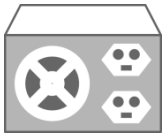
LSLA: MoA, not MoHW

- LSLA models HW + communication libraries + scheduler + Oss +…
- LSLA models the service the platform offers to the applications
- Top-down approach
  - **Learning parameters from experiments**

System Objectives

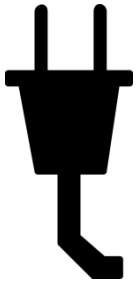


Reliability



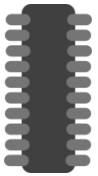
Peak Power

T°C



Energy

Latency



Memory

Security

$

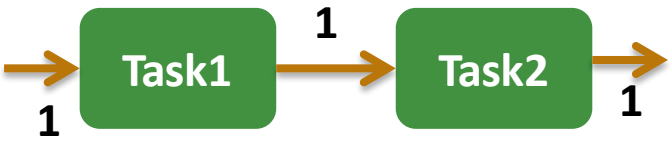
Unit Cost

$

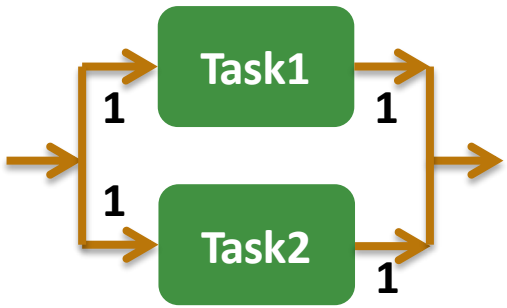
Maintenance Cost

MoAs: Limits of LSLA

- Energy  ← Linear model OK
- Latency 
- Latency does not have an additive nature

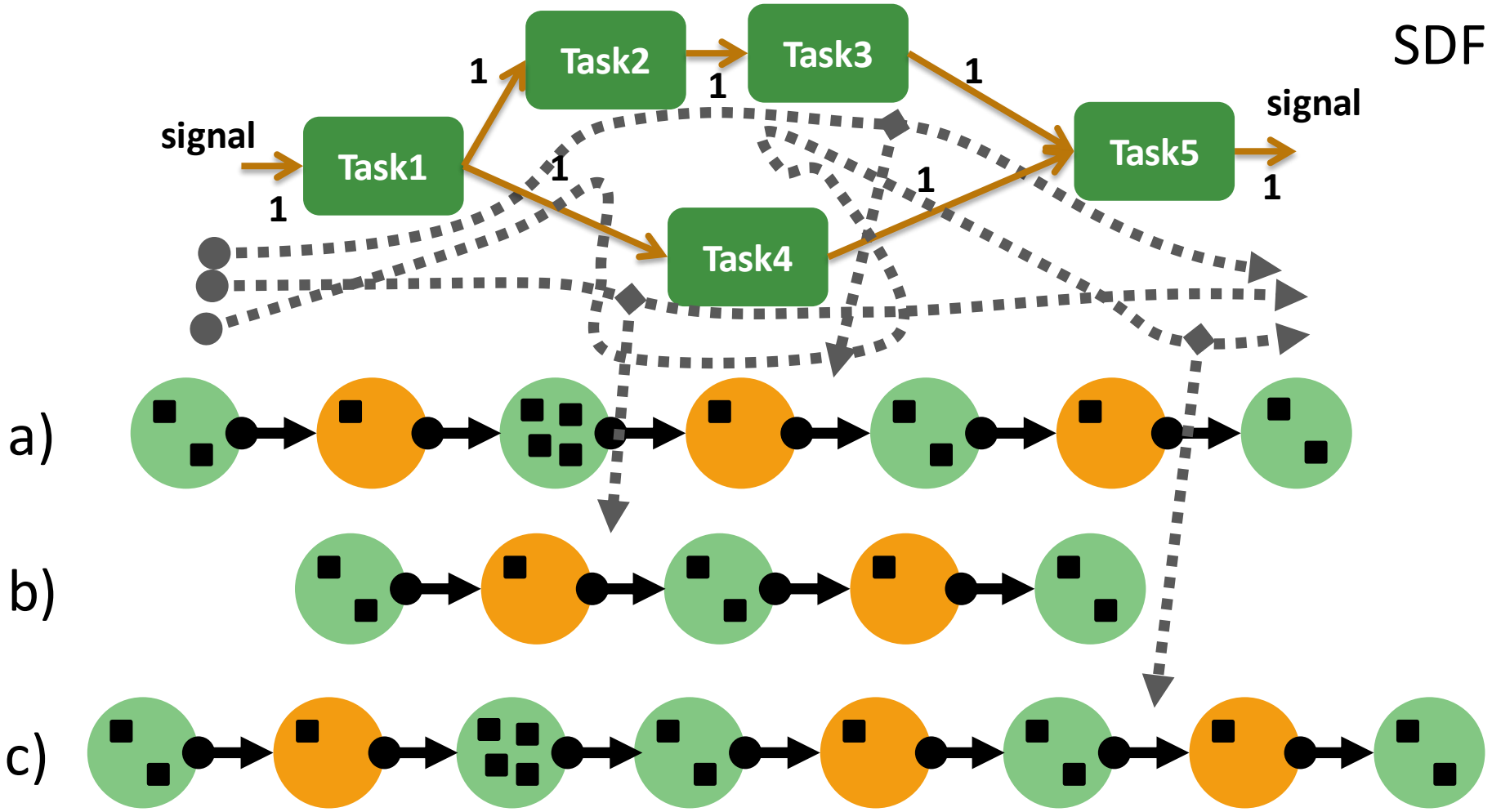


→ **Latency = sum**

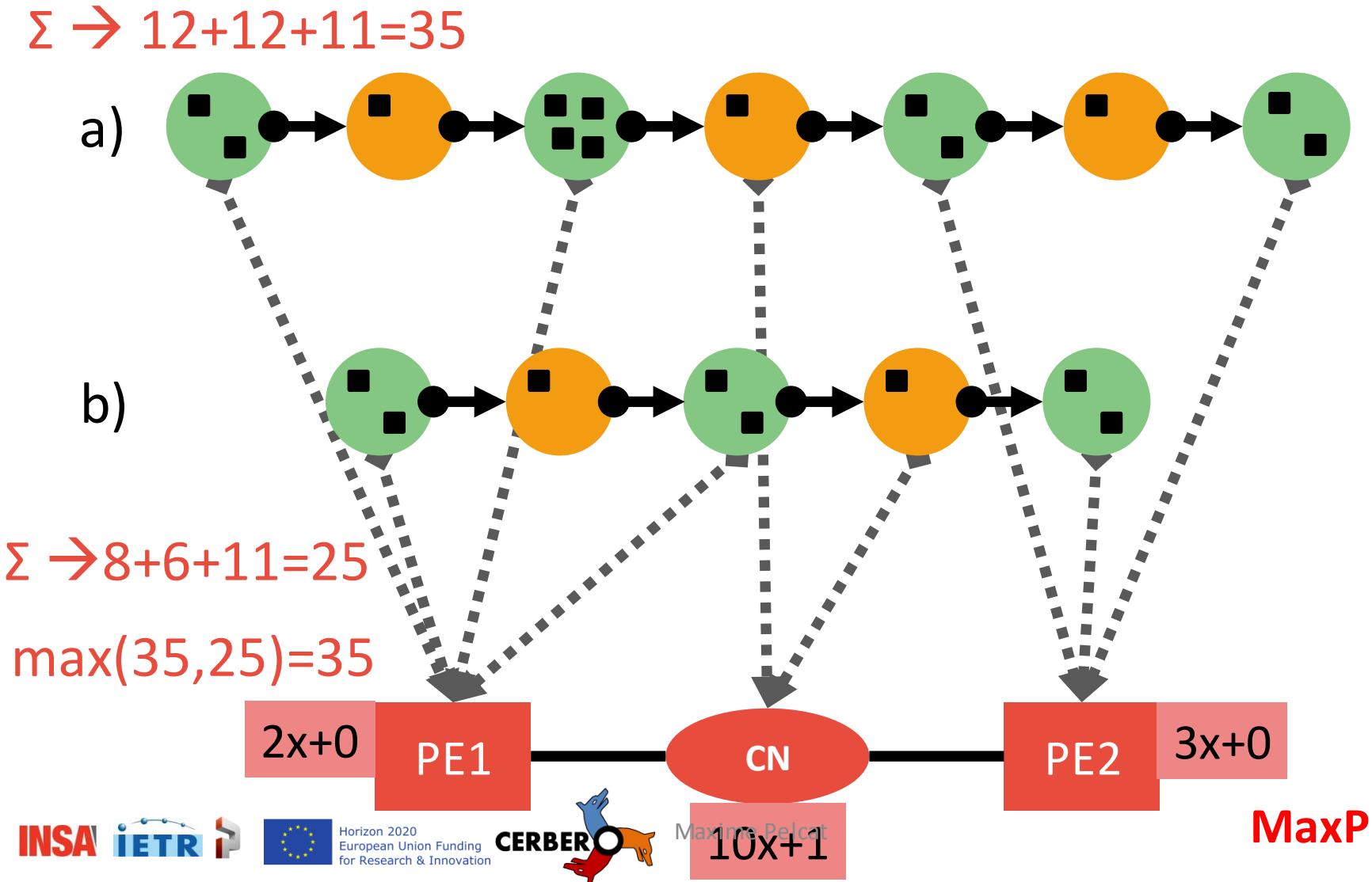


→ **Latency = max**

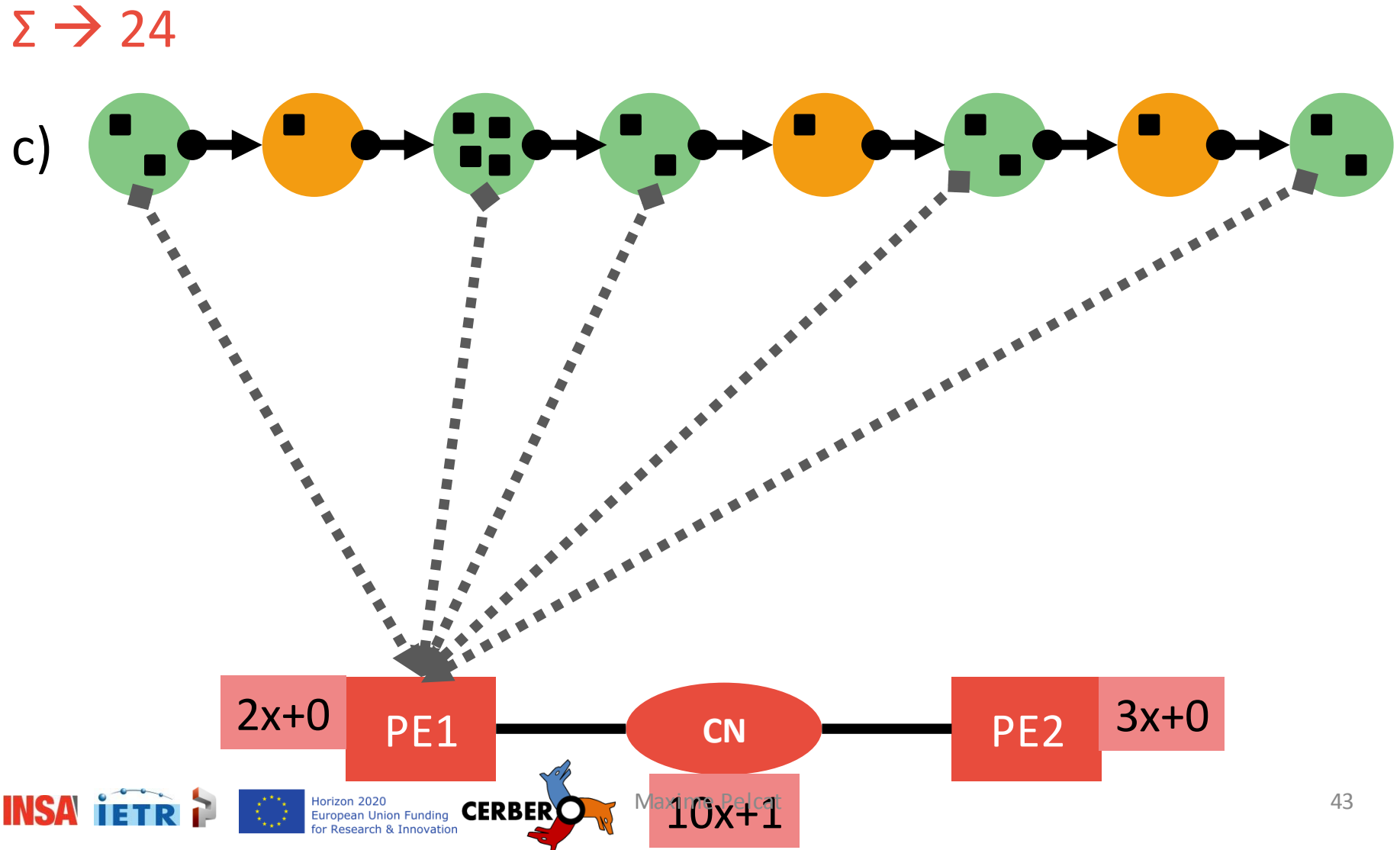
Activity & MoA for Latency



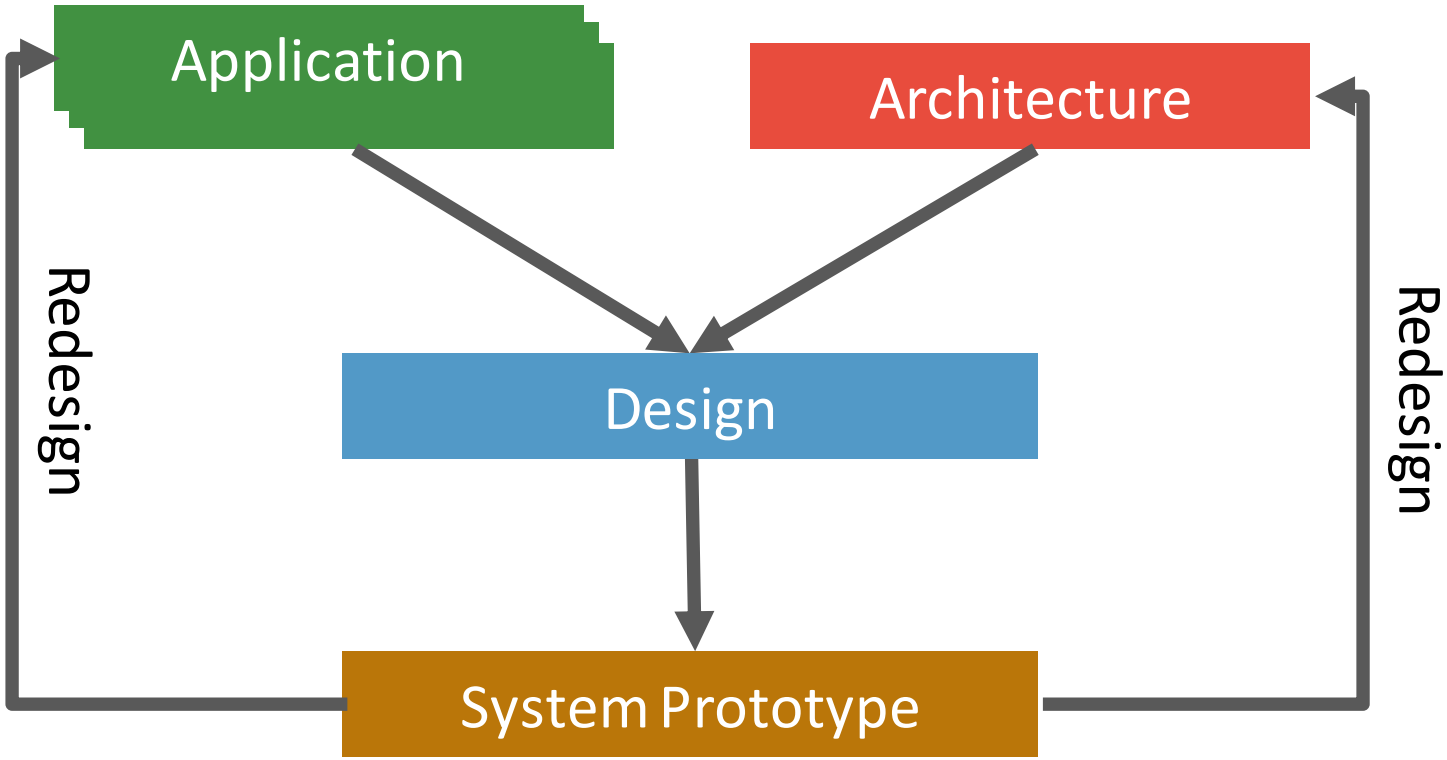
Activity & MoA for Latency



Activity & MoA for Latency



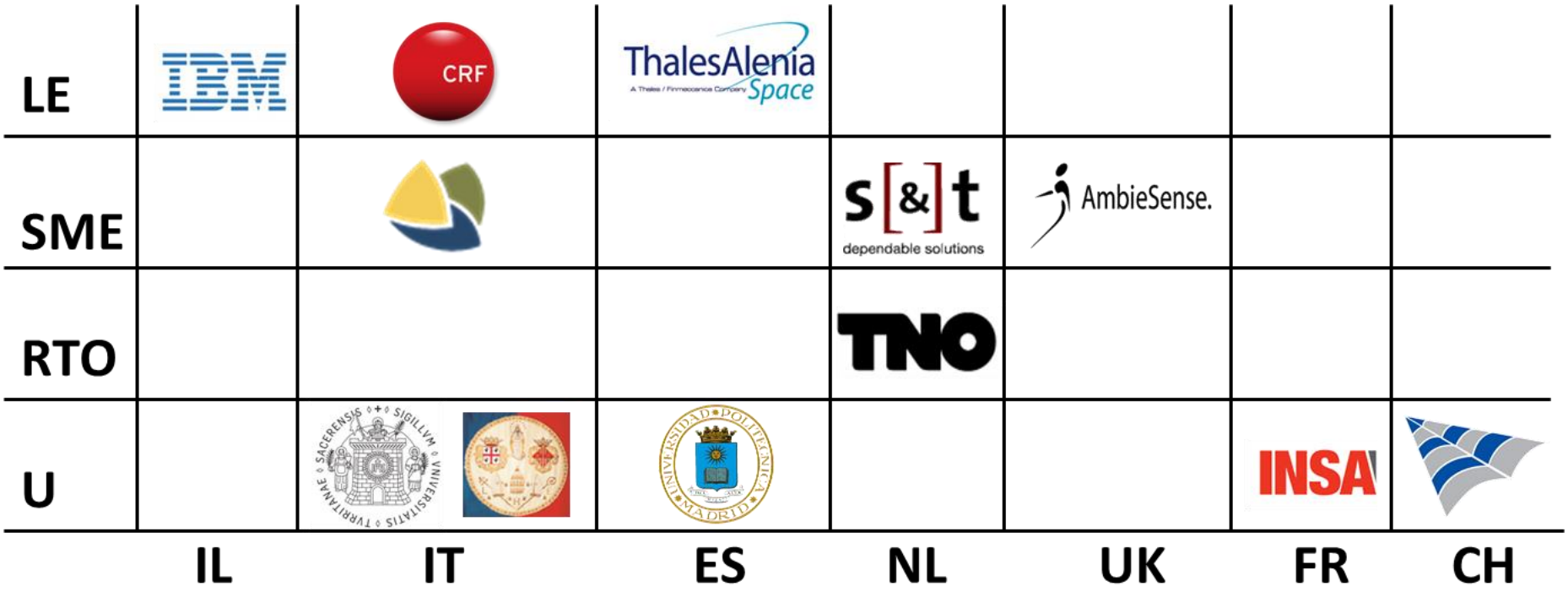
Accuracy? No, Fidelity!



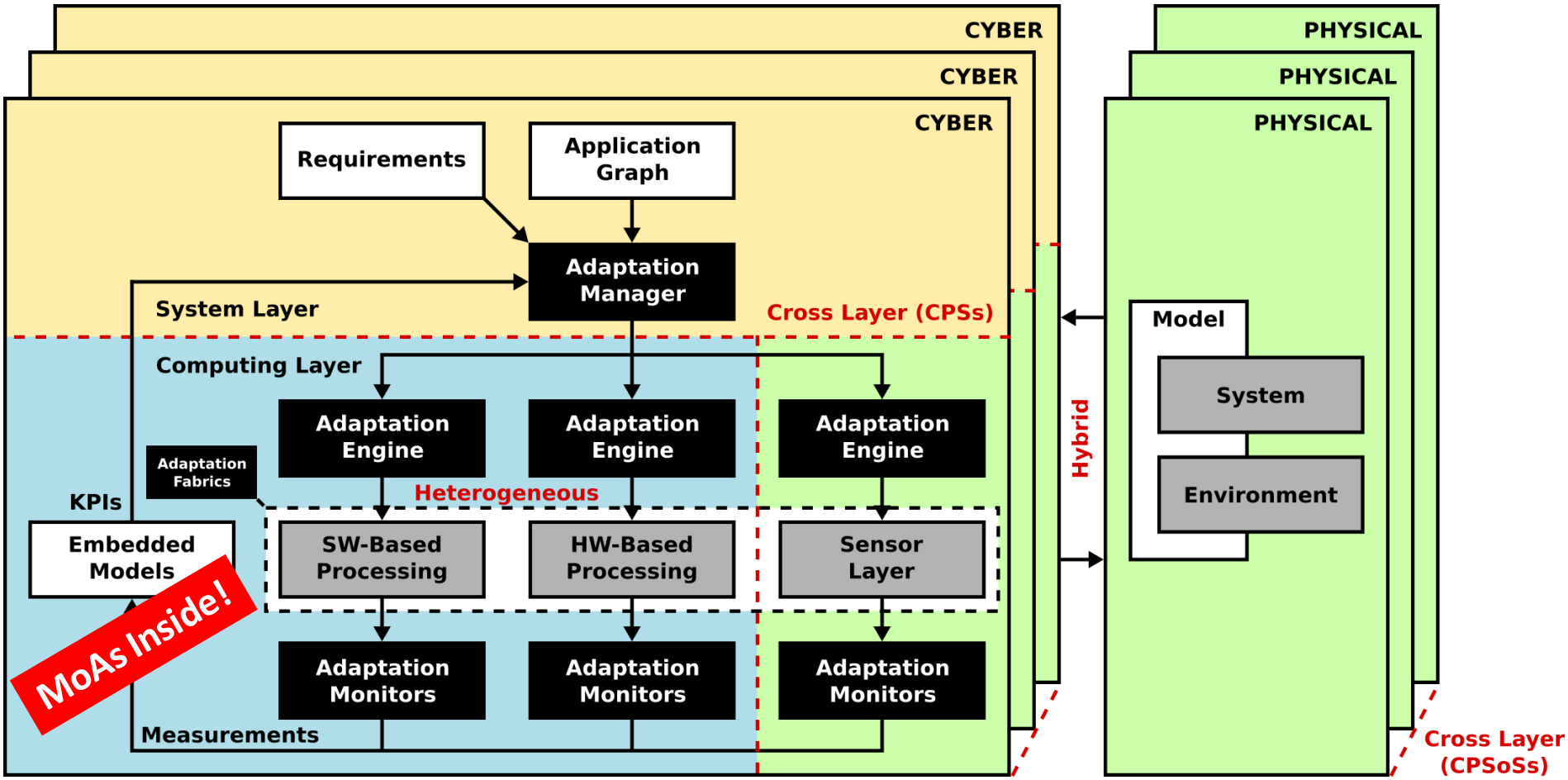
Current Activities

H2020 CERBERO

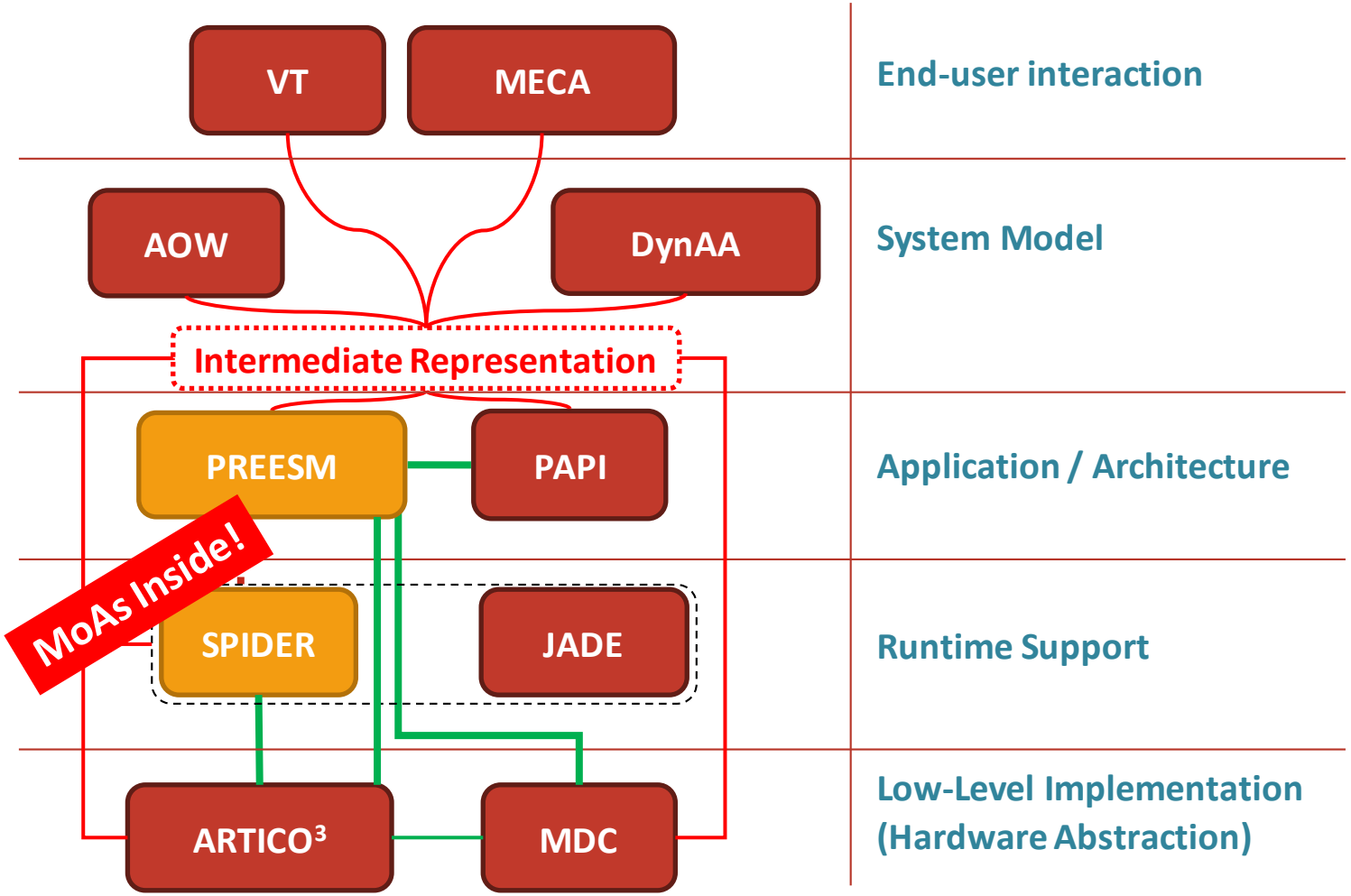
- **Cross-layer Design** of **Reconfigurable Cyber-Physical Systems**



CERBERO System Adaptation



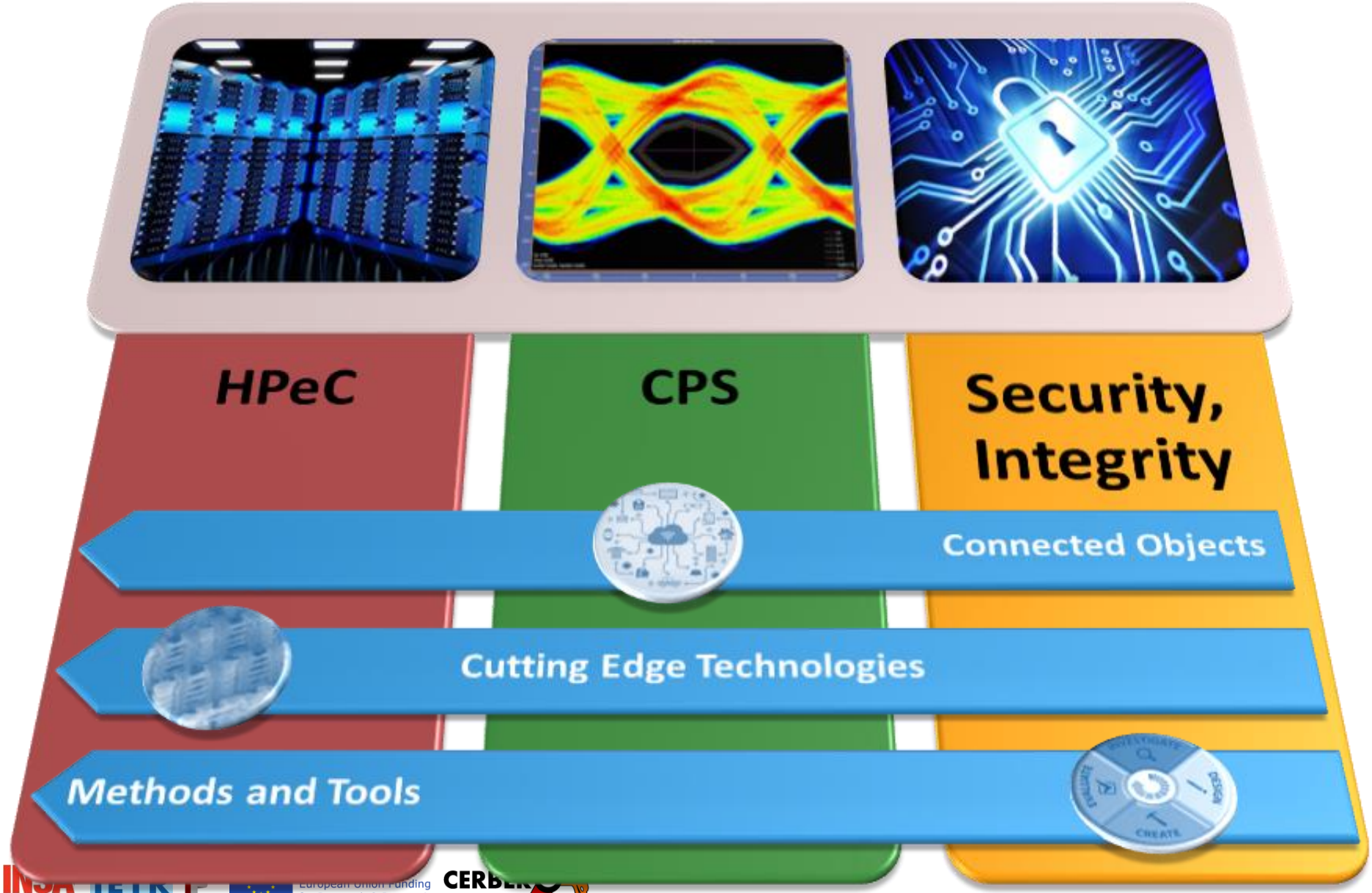
H2020 Cerbero Toolchain



GdR SOC2

- Groupement de recherche SOC2
  - **Systems on a Chip, Connected Systems**
  - **Industrial partner club**

GdR SOC2


HPeC
CPS
Security, Integrity
Connected Objects
Cutting Edge Technologies
Methods and Tools

SAMOS XIX

- 19th edition of SAMOS Conference
- July 7-11, submission in March

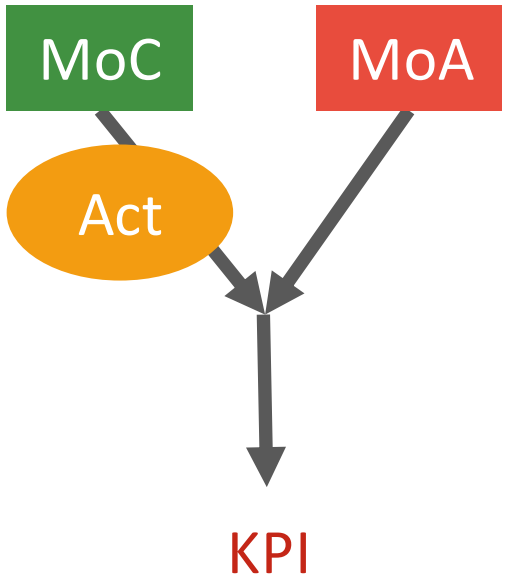
International Conference on Embedded Computer Systems: Architectures, MOdeling and Simulation

SAMOS XIX

Pythagoreio, Samos Island, Greece
July 7 – 11, 2019

Takeaway Message

- MoAs can early predict performance/quality
  - **Especially for HPeC systems**
- MoAs are not HW Models
  - **They model HW + protocols + OS + …**
- MoAs are built/learnt top-down
  - **They can and should be simple**
- The need for MoAs may rise
  - **Due to Fog/Edge Computing complexity**



Questions?

PREESM

http://preesm.org

International Conference on Embedded Computer Systems: Architectures, MOdeling and Simulation

SAMOS XIX

Pythagoreio, Samos Island, Greece
July 7 – 11, 2019

www.cerbero-h2020.eu

Pelcat, M; Mercat, A; Desnos, K; Maggiani, L; Liu, Y; Heulot, J; Nezan, J-F; Hamidouche, W; Ménard, D; Bhattacharyya, S (2017) "Reproducible Evaluation of System Efficiency with a Model of Architecture: From Theory to Practice", IEEE TCAD.