PRIN COMMON-WEARS

Ongoing research (@UNITO)

(December 19, 2022)

Ferruccio Damiani (UNITO)

People @ UNITO & C

@UNITO

- **Giorgio Audrito** (RTD B)
- **Ferruccio Damiani** (PO)
- **Gianluca Torta** (RU)
- **Daniele Bortoluzzi** (Neolaureato, supervisionato da Gianluca Torta e Andrea Basso, vincitore Assegno di Ricerca, 2 anni a partire da marzo 2023), developer
- Lorenzo Testa (Dottorando, ciclo 36, Alto Apprendistato in Concept Reply)
- Marco Ottina (Dottorando, ciclo 37, Dottorato nazionale in AI - Industry 4.0)
- Yasir Shabir (Dottorando, ciclo 37, PON Green - con Synesthesia)

@Syneshesia (https://synesthesia.it/)

- **Andrea Basso** (https://www.linkedin.com/in/abasso1/): IoT/Edge AI, IoT security, Standardization,...

Aggregate Programming (AP)

Formal Methods and **Rigorous Software Engineering** perspective (when feasible)

ONGOING

1. FCPP Extensions [presented by Giorgio Audrito]
2. Secure FCPP [presented by Gianluca Torta]
3. AP for Cooperative ML Inference [presented by Gianluca Torta]

PLANNED

4. Testing for AP
5. Variability modeling for AP

Language Design, Algorithms and Properties of AP

Ongoing work:

- Dynamics and applications of aggregate processes
- Adding predictive capabilities to runtime verification

Planned work:

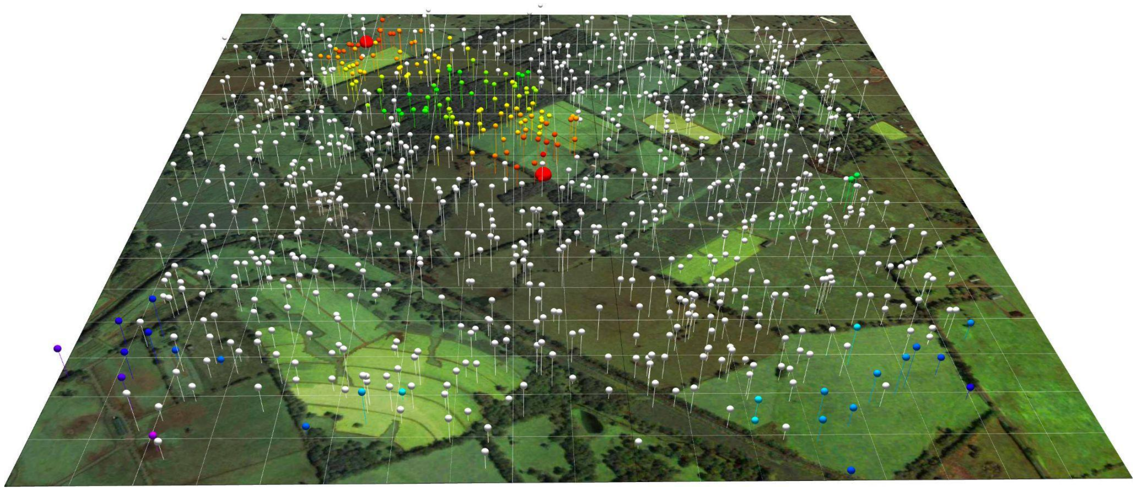
- Real-time guarantees for state-of-the-art resilient algorithms

FCPP Extension

FCPP tool chain [presented by Giorgio Audrito]

https://fcpp.github.io/

- efficient, portable, small footprint
- integrated graphical simulator for development and debugging
- **designed for execution on multiple platforms**



FCPP platform support

Some initial support:

- WandStem microcontrollers running MIOSIX (with PoliMI)
- Smartphones running Android (with HVL)
- HPC architectures for hybrid cloud-edge systems

Planned work:

- Linux companion computers in embedded systems
  - …also for **drones**

Industrial application scenario (with Synesthesia): swarms of drones and rovers

In the context of the CN AGRITECH we are buying

- a fleet of UAVs, two kinds
    - Crazyflie 2.1(https://www.bitcraze.io/products/crazyflie-2-1/), weight: 27g
    - **DJI Matrice M300 RTK**, weight: ~6.3Kg, payload: 2.7Kg
- budget: 50K Euro

In the context of the PNRR NODES Spoke1 (Mobility) we are:

- buying fleet of small rovers
- integrating with fleet of Crazyflies
- developing a demo based on logistics

Engineering Tools

There isn’t much existing literature on methods addressing engineering tasks such as analysis, design, development, deployment and testing on AP systems.

- An initial study of the possible directions to explore can be found in:
  Casadei R, Pianini D, Aguzzi G, Audrito G, Torta G, Ottina M, Damiani F, Viroli M. *Towards Automated Engineering for Collective Adaptive Systems: Vision and Research Directions.* Proc. 1st International Workshop on COMMunity-OrieNted WEARrable Computing Systems, IEEE 2022

- Planned work on automated testing of aggregate programs in FCPP:
  genetic algorithms for generating critical simulations giving rise to errors

IoT Security and FCPP

Security on Embedded Systems

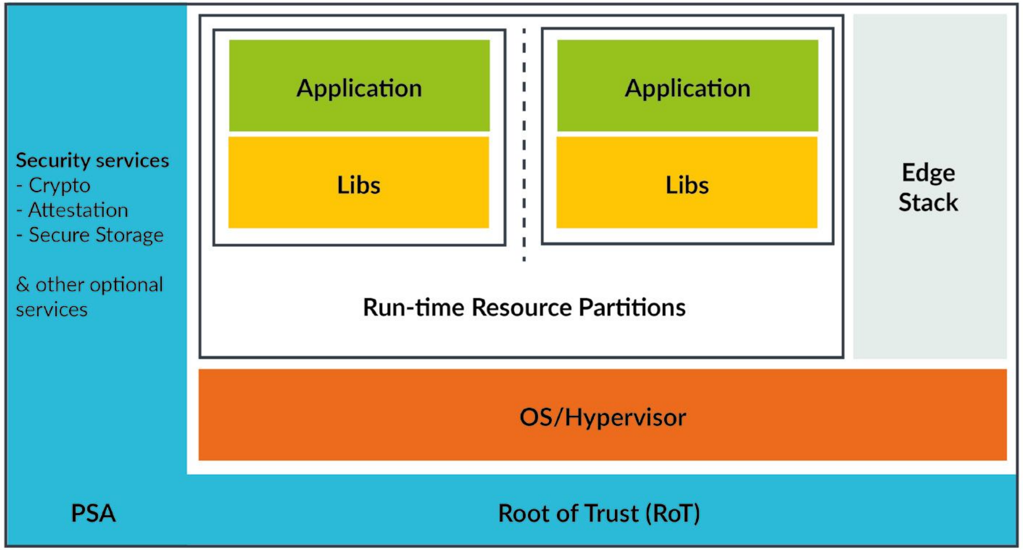
- several factors currently making the **cost of security** a concern for the **silicon vendors**, **software providers** and **device manufacturers** vital to IoT success [1]
- as IoT adoption grows, so does the **cyber risk** from bad actors wanting to exploit it
  - it's estimated that by 2025 cybercrime damages will total over $10 trillions
- PSA (Platform Security Architecture) [2]
  - security framework that allows security to be consistently designed in, at both a **hardware** and **firmware** level
  - fourth and final stage is **PSA Certified**, which currently offers certification for constrained IoT devices via an independent body

[1] PSA Certified. The IoT Industry Action Plan to Reduce the Cost of Security. White Paper (2022)

[2] ARM. The Importance of Security for the Infrastructure Edge. White Paper (2019)

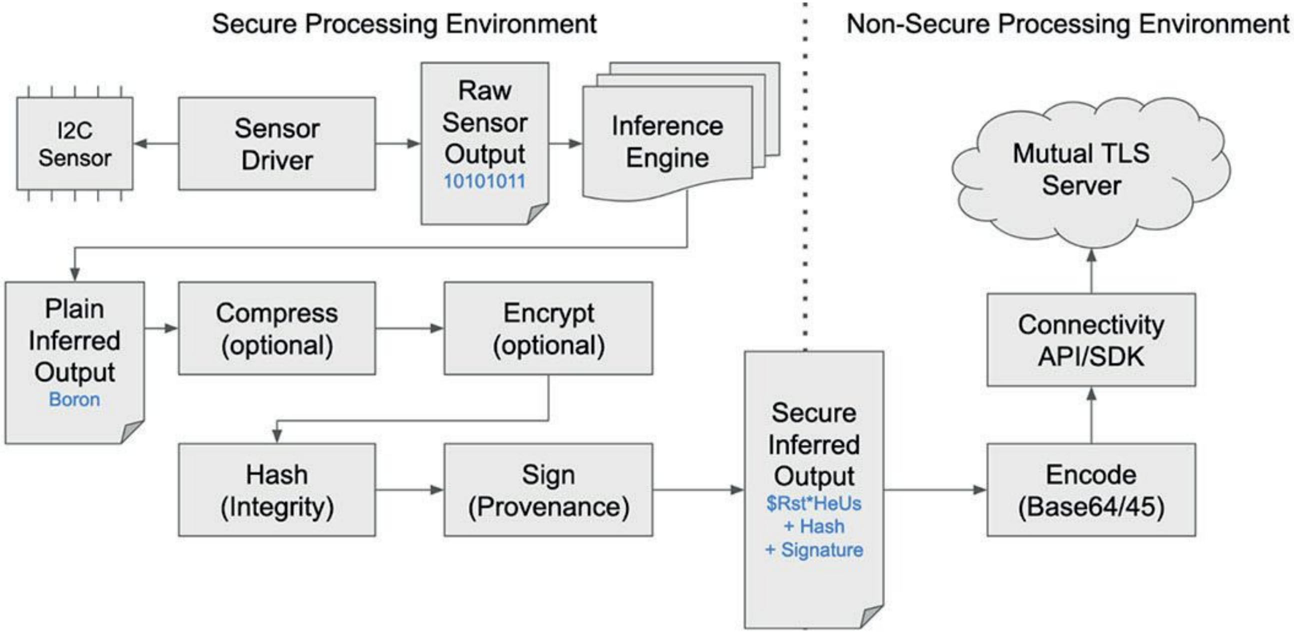
Secure HW and Firmware

- separate the world into “secure” (S) and “non-secure” (NS) at HW level (**ARM TrustZone**)
- security services in S partition (**Trusted Firmware M**)
    - crypto, attestation, secure storage
- further isolate (if needed) applications from each other



Secure ML

- Linaro is defining a framework for secure ML on MCUs (Confidential AI [1])



[1] B. Fletcher (Linaro). Confidential AI for MCUs. White Paper (2021)

Secure FCPP

- port FCPP to **STM32U585AI** MCU → Arm Cortex-M33 core with **TrustZone** and Armv8-M security extension
- ZephyrOS
  - fully supports **TF-M integration**
  - supports **more than 400 boards**
- TF-M security services
  - proximity-based **communication** → encryption/decryption, authentication, …
  - isolate AP **applications** running on top FCPP
- analogies with V2V communication
  - VPKI (public key infrastructure)
  - problem of scale [1]

[1] M. A.Simplicio , E. L. Cominetti , H. K. Patil, J. E. Ricardini and M. V. M. Silva, "The Unified Butterfly Effect: Efficient Security Credential Management System for Vehicular Communications," 2018 IEEE Vehicular Networking Conference (VNC) VNC), 2018

AP for Cooperative ML Inference

ML Inference at the (Far) Edge

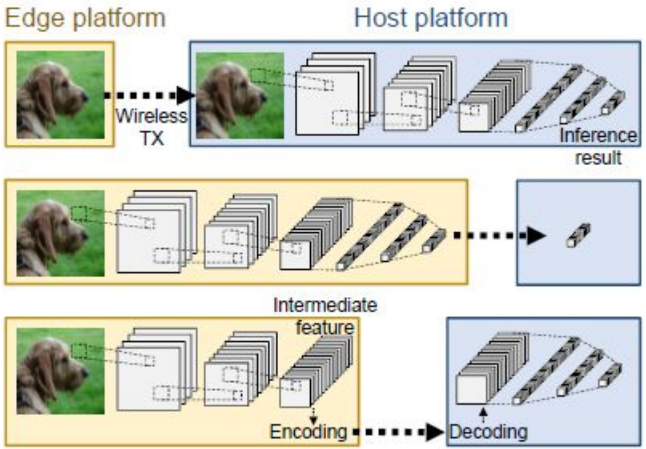
- performing (part of) **DNN inference on small constrained devices** at the far-edge is increasingly important [1,2]
  - privacy
  - latency
  - communication/energy costs
- **quantization** and **pruning** are well known techniques
- when devices are very low-end (TinyML) also **Neural Network partitioning** and **cooperative inference**

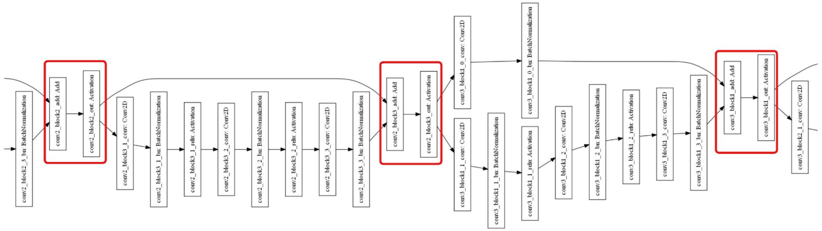
[1] Martins Campos de Oliveira F, Borin E. Partitioning convolutional neural networks to maximize the inference rate on constrained IoT devices. Future Internet. 2019 Sep 29;11(10):209

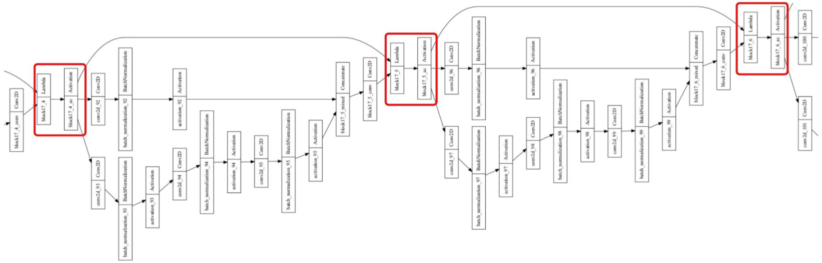
[2] Ko JH, Na T, Amir MF, Mukhopadhyay S. Edge-host partitioning of deep neural networks with feature space encoding for resource-constrained internet-of-things platforms. In2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS) 2018 Nov 27 (pp. 1-6). IEEE.

Partition and Cooperative Inference

- partition computed in **centralized** or **distributed** way
- generates a **directed acyclic graph** of computation
- in **AP** proximity-based device network topology
    - **partition DAG** mapped to **device network**

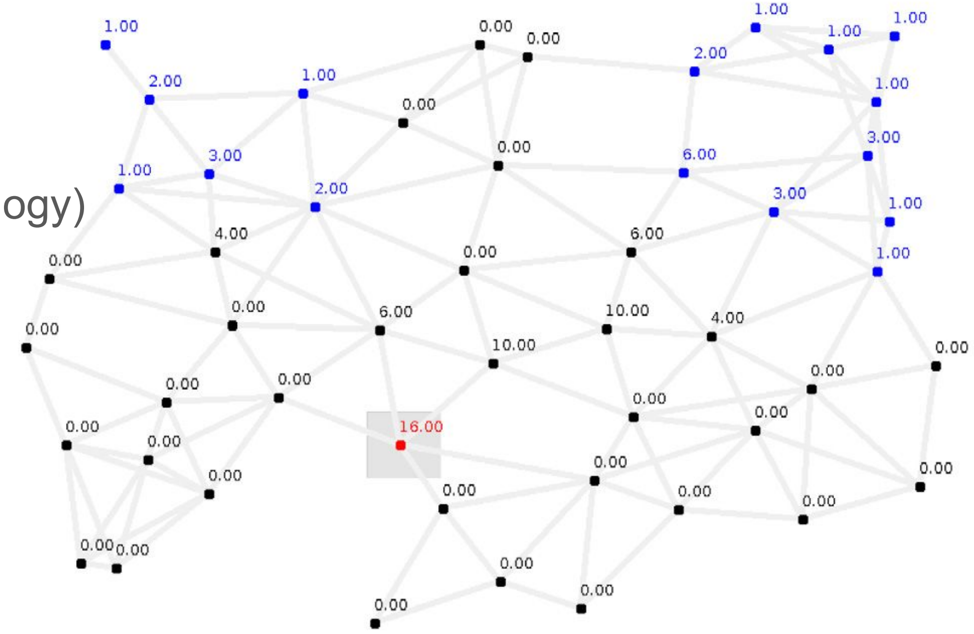



conv2_block2_3_bn: BatchNormalization
conv2_block2_add: Add
conv2_block2_out: Activation
conv2_block3_1_conv: Conv2D
conv2_block3_1_bn: BatchNormalization
conv2_block3_1_relu: Activation
conv2_block3_2_conv: Conv2D
conv2_block3_2_bn: BatchNormalization
conv2_block3_2_relu: Activation
conv2_block3_3_conv: Conv2D
conv2_block3_3_bn: BatchNormalization
conv2_block3_add: Add
conv2_block3_out: Activation
conv3_block1_0_conv: Conv2D
conv3_block1_0_bn: BatchNormalization
conv3_block1_1_conv: Conv2D
conv3_block1_1_bn: BatchNormalization
conv3_block1_1_relu: Activation
conv3_block1_2_conv: Conv2D
conv3_block1_2_bn: BatchNormalization
conv3_block1_2_relu: Activation
conv3_block1_3_conv: Conv2D
conv3_block1_3_bn: BatchNormalization
conv3_block1_add: Add
conv3_block1_out: Activation
conv3_block2_1_conv: Conv2D



Adaptivity and Resilience with AP

- AP supports computation on a DAG
  - “collection” of result
- self-adapt to:
  - additional computational nodes
  - nodes failure/vanishing
  - nodes movements (change of topology)
  - improve performance


1.00
1.00
1.00
0.00
1.00
0.00
2.00
1.00
2.00
1.00
0.00
3.00
0.00
3.00
6.00
1.00
2.00
3.00
1.00
4.00
6.00
0.00
1.00
0.00
0.00
6.00
10.00
4.00
0.00
10.00
0.00
0.00
0.00
0.00
0.00
0.00
0.00
16.00
0.00
0.00
0.00
0.00
0.00
0.00
0.00
0.00
0.00
0.00
0.00
0.00