Neural Networks on STM32

Artificial Intelligence Solutions

FUTURE ELECTRONICS

Practical Machine Learning and AI Seminars

life.augmented

FUTURE ELECTRONICS

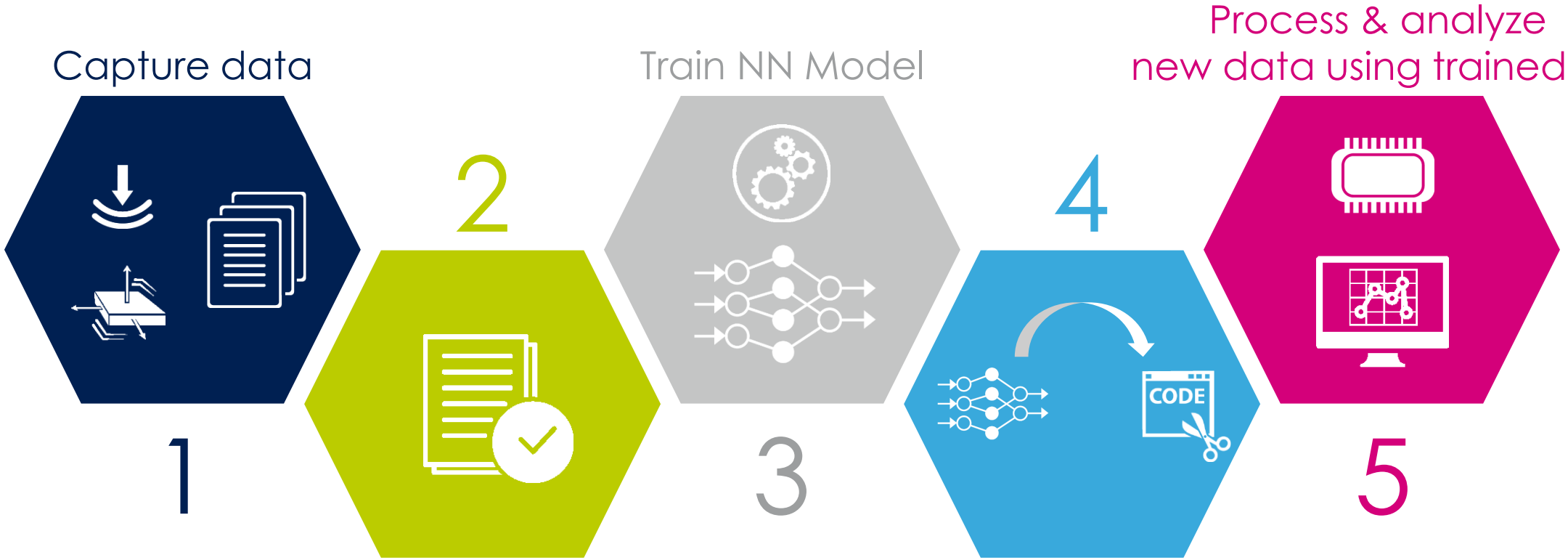
Neural Networks on STM32

Simple, fast, optimized

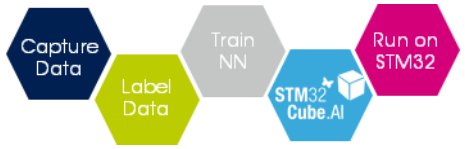
STM32 Cube.AI

The Key Steps Behind Neural Networks

Neural Network (NN) Model Creation | Operating Mode



life.augmented



ST Toolbox for Neural Networks

life.augmented

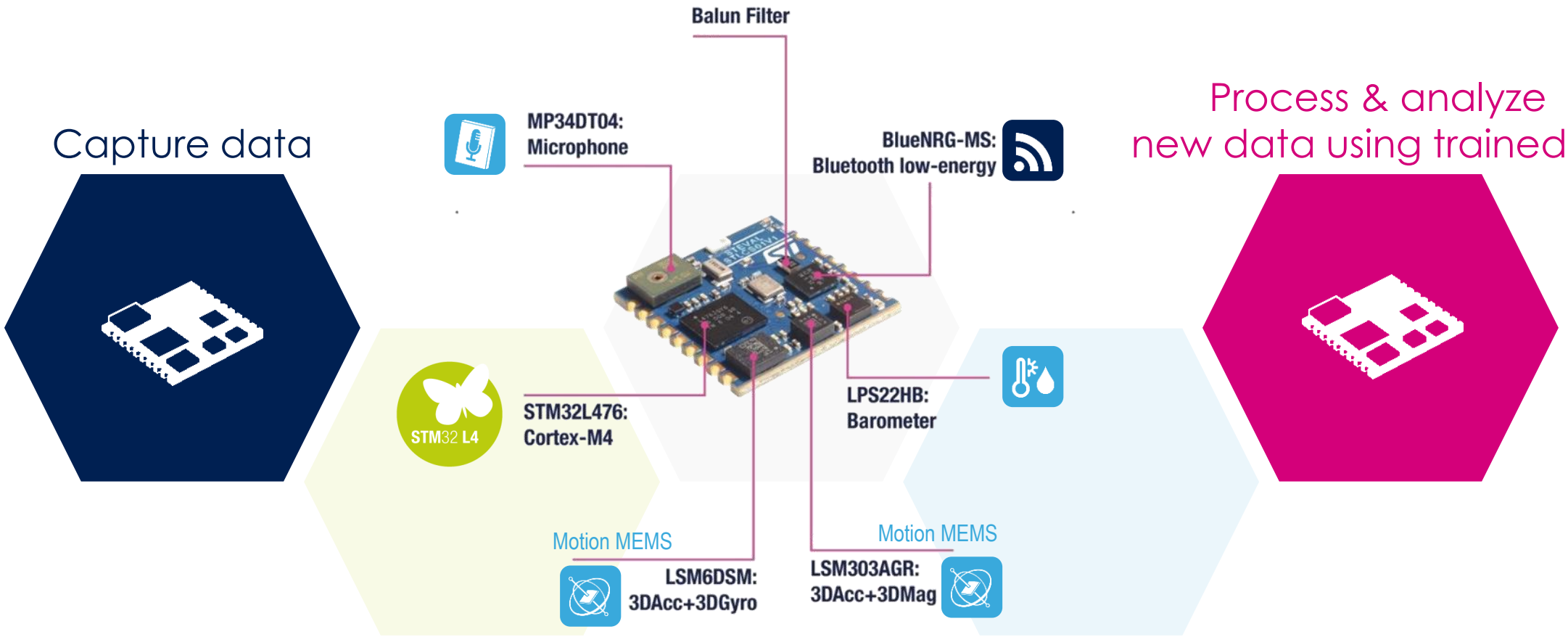
Capture data

Clean, label Data
Build NN topology

Convert NN into
optimized code for MCU

Process & analyze
new data using trained NN

Form Factor Hardware
to Capture and Process Data



www.st.com/SensorTile
www.st.com/SensorTile-edu

life.augmented

Form Factor Hardware

AI IoT Node for More Connectivity

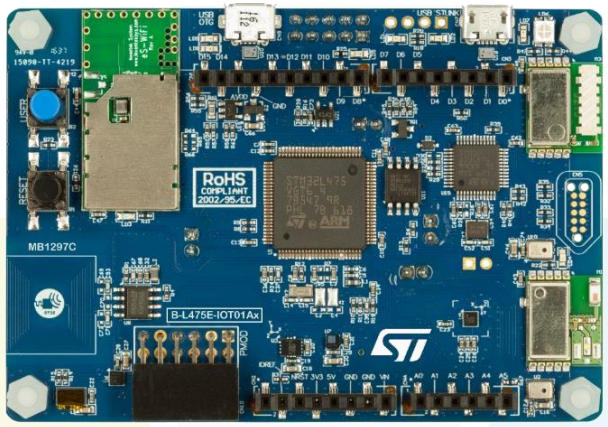
Capture Data | Label Data | Train NN | STM32 Cube.AI | Run on STM32

Sub 1GHz Sub-1GHz

NFC Dynamic NFC Tag

Wi-Fi Wi-Fi

Capture data



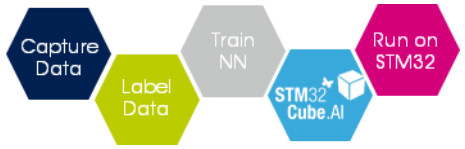
Process & analyze new data using trained NN

More debug capabilities

- Integrated ST-Link/V2.1
- PMOD extension connector
- Arduino Uno extension connectors

https://www.st.com/iotnode

life.augmented



Collecting Data & Architecting a NN Topology

Services provided by Partners

Capture data

Clean, label Data
Build NN topology

ST tools to support

ST BLE Sensor

ST BLE Sensor mobile phone application

Collect and label data from the SensorTile

life.augmented

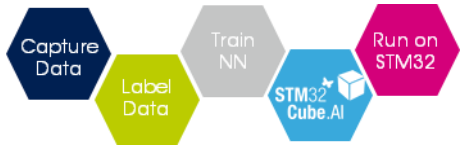
Partner Program

Selected partners

Neural Networks engineering services support

Data scientists and Neural network architects

life.augmented



STM32CubeMx Extension

More Than an AI Conversion Tool

Input your Framework dependent, Pre-Trained neural network into **STM32Cube**.AI

Automatic and fast generation of an STM32-optimized library

STM32Cube.AI guarantees interoperability with state-of-the-art Deep Learning Design Frameworks

Train NN Model

ConvNetJS
Deep Learning in your browser

Lasagne

Caffe

* TensorFlow used as a Keras backend. Not all operators accessible to MCUs

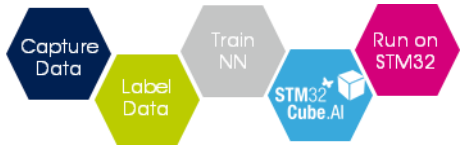
STM32 Cube.AI

Convert NN into optimized code for MCU

Process & analyze new data using trained NN

STM32

life.augmented



STM32Cube.AI Roadmap

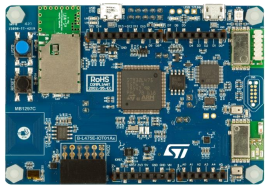
2019

- Floating Point Support

Keras
ConvNetJS
Deep Learning in your browser
Lasagne
Caffe

Apr

IoTNode



Jun

- Quantization
- TensorFlow for MCU
- Command line interface
- UI Improvements
- Additional layers

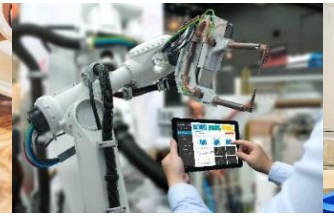
+ TensorFlow Lite microcontrollers

2020

- ONNX Introduction
- Additional layers
- Debug improvements

Keras + ONNX mxnet Microsoft CNTK Chainer PyTorch



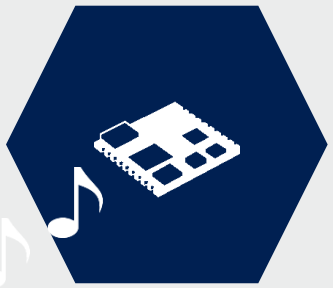





life.augmented

Audio Scene Classification (ASC)

Audio Example in FP-AI-SENSING1 Package



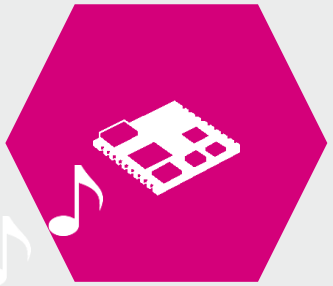
3 classes

Embedded **audio** pre-processing

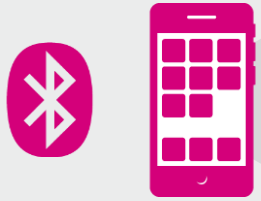
Data sent to mobile app for **labelling**

Data stored on the device SD card for future **labelling**

Indoor, Outdoor, In vehicle



NN & example dataset provided



Embedded **audio** pre-processing

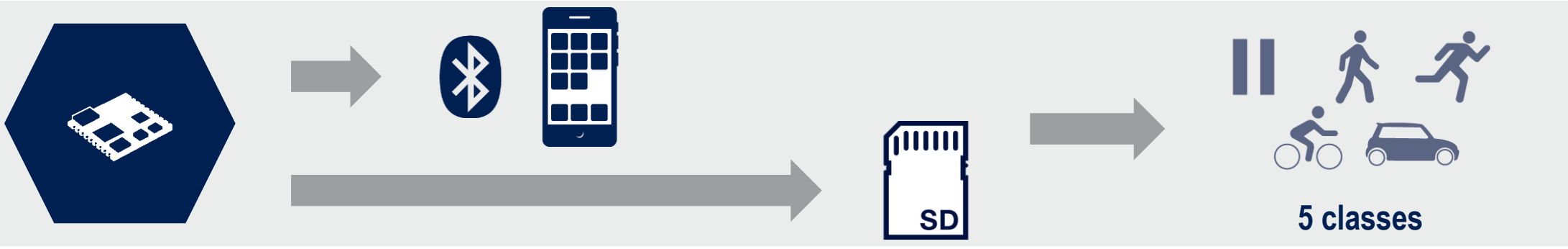
Inferences running on the microcontroller

Inference result displayed on mobile app

life.augmented

Human Activity Recognition (HAR)

Motion Example in FP-AI-SENSING1 Package



Embedded **motion** pre-processing

Data sent to mobile app for **labelling**

Data stored on the device SD card for future **labelling**

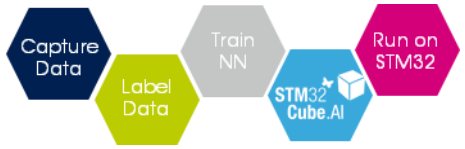
Stationary, walking, running, biking, driving

Embedded **motion** pre-processing

Inferences running on the microcontroller

Inference result displayed on mobile app

life.augmented



STM32 Solutions for AI

More Than Just an STM32CubeMx Extension

An extensive toolbox to create your AI application easily

AI extension for STM32CubeMX

To map pre-trained Neural Networks

STM32 Cube.AI

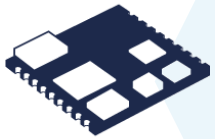
Resources

Function packs for Quick prototyping

Audio and Motion examples

SensorTile reference hardware

To run inferences or data collection



… And more coming!

STM32 Community with dedicated Neural Networks topic

Mobile phone application

To collect and label data

To display the result of inference running on the STM32

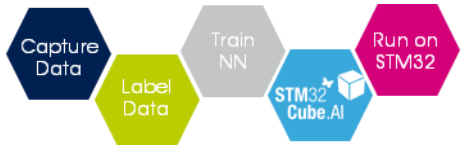
life.augmented

Partner Program

ST Partner Program with dedicated group for Partners providing Neural Networks engineering services

Data scientists and Neural network architects

life.augmented



Making AI Accessible Now

Leader in Arm® Cortex®-M 32-bit General Purpose MCU

Compatible with STM32Cube.AI ecosystem

Compatible with Partner Machine Learning ecosystems

2007 — World 1st Cortex-M MCU — STM32 F1

2009 — World 1st Cortex-M Ultra-low-power — STM32 L1

2010 — 1st High Perf. 120 MHz, 90nm — STM32 F2

2011 — 1st High Perf. Cortex-M4 168 MHz — STM32 F4

2012 — 1st Mixed Signal DSP + Analog STM32F3 Cortex-M4 — STM32 F3

2012 — Entry Cost STM32F0 Cortex-M0 — STM32 F0

2013 — Entry Cost Ultra-low-power — STM32 L0

2014 — World 1st Cortex-M7 — STM32 F7

2015 — Leadership Ultra-low-power Cortex-M4 — STM32 L4

2015 — 273 ULPBench™ #1 ULP

2016 — #1 Performance 2020 CoreMark — STM32 H7

2016 — Ultra-low-power Excellence — STM32 L4+

2017 — Mainstream Cortex-M0+ MCUs Efficiency at its best! — STM32 G0

2018 — Dual-core, multi-protocol and open radio — STM32 WB

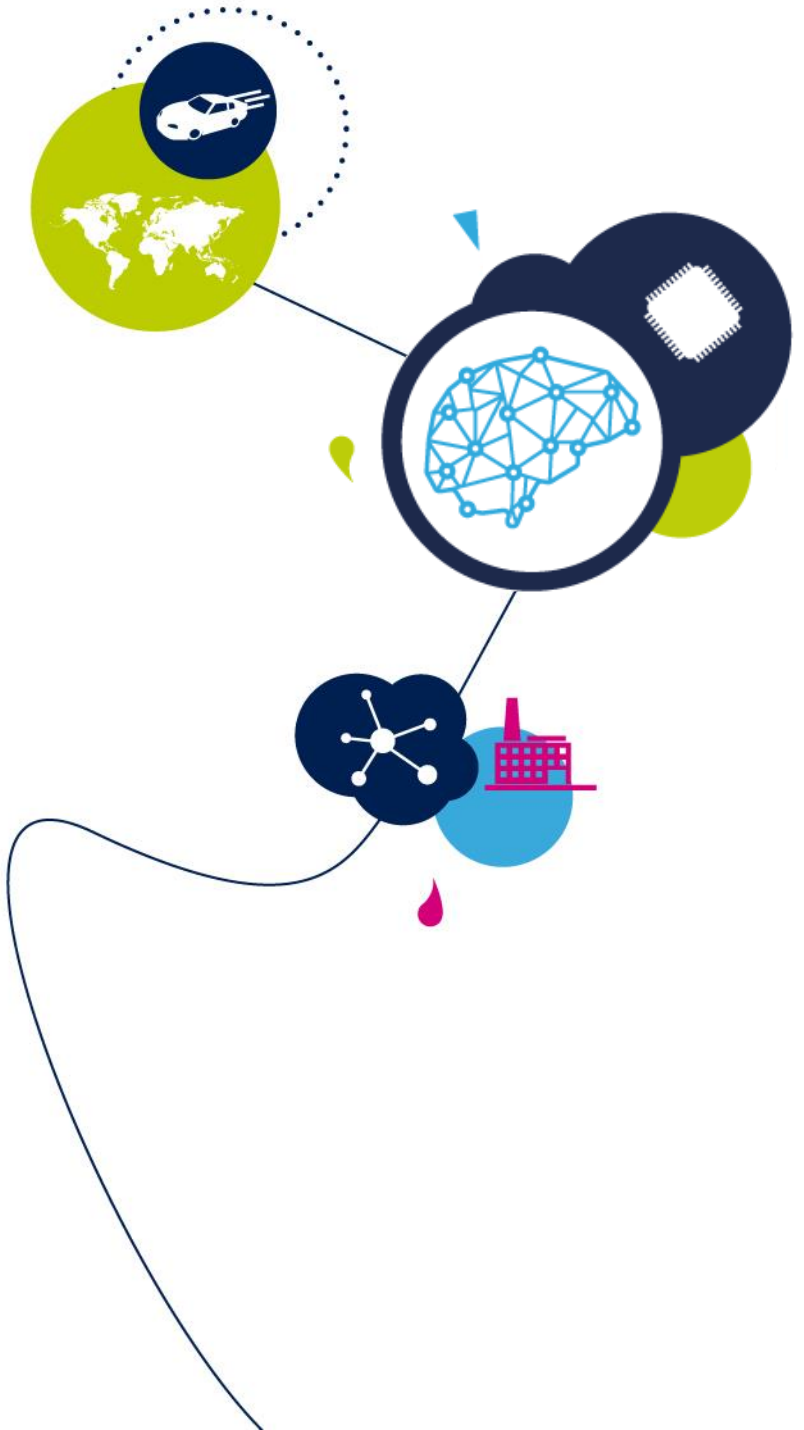
2019 — First STM32 MPU Dual Cortex-A7 + Cortex-M4 STM32 meets Linux — STM32 MP1

2019 — Introduction of M33 Excellence in ULP with more security — STM32 L5

life.augmented

More than 40,000 customers

Over 4 Billion STM32 shipped since 2007

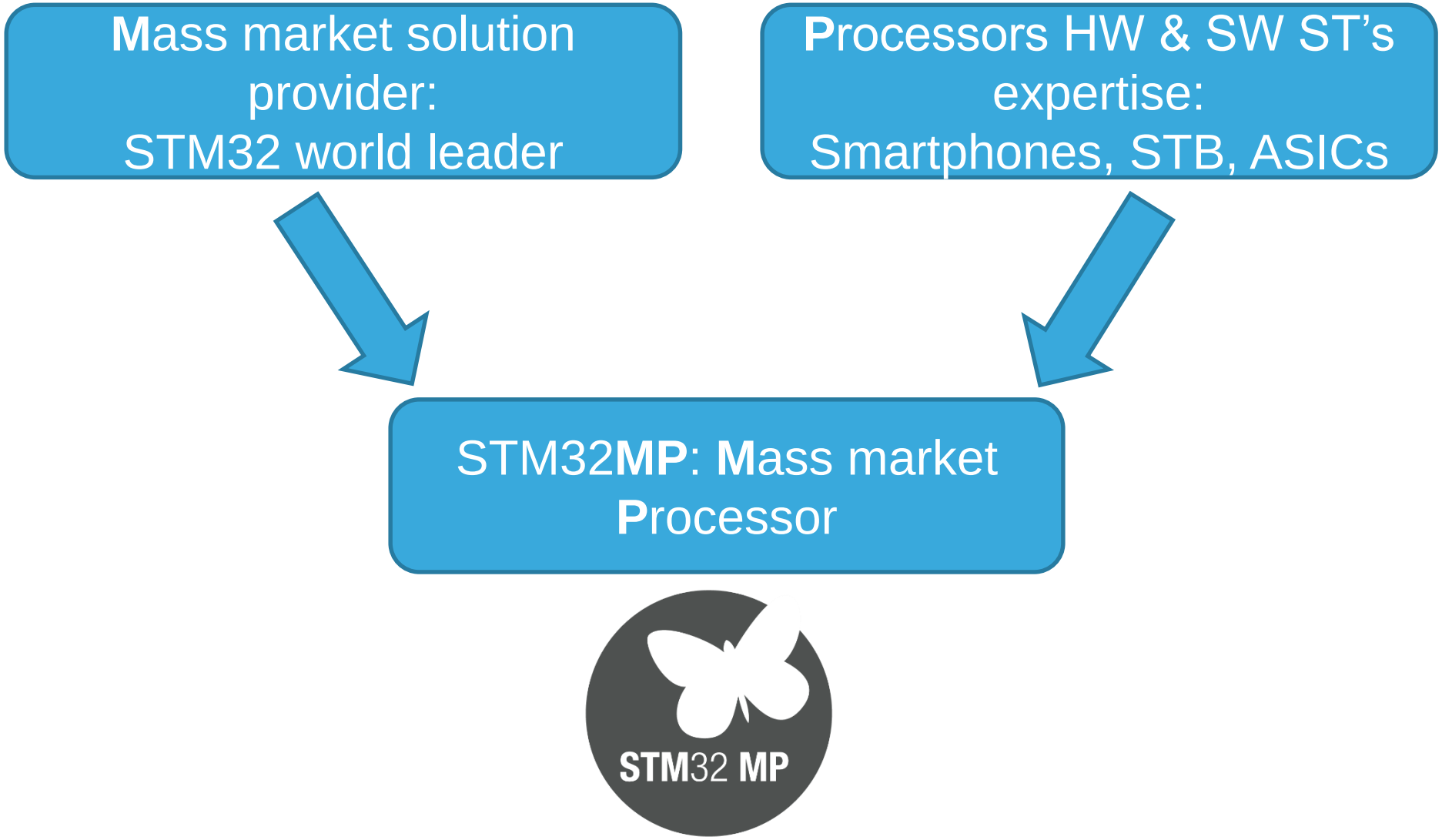


STM32MP1 Presentation

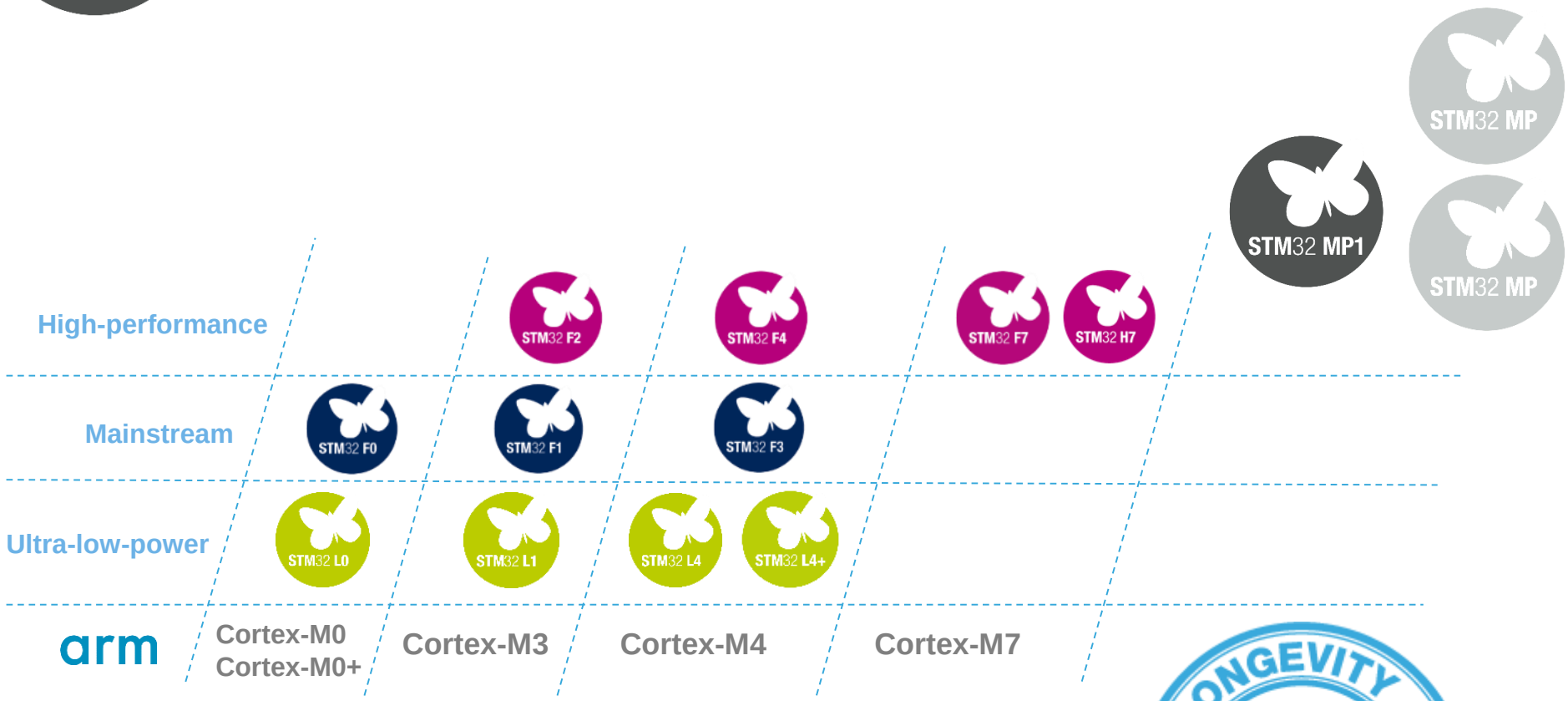
STMicroelectronics

life.augmented

What is STM32MP ?



Embedded Processing Continuum



life.augmented

STM32MP1 Product Positioning

- Roadmap extension of STM32H7 and 1st ST's Mass Market MPU
- Low to Mid range MPU class

STM32MP1

Arm dual Cortex-A7

3D GPU

DRAM Interface

Linux Ecosystem

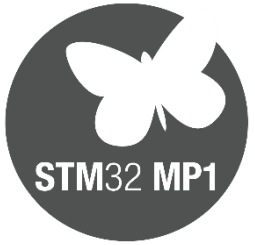
STM32 MCUs Reuse

Arm Cortex-M4

STM32 IPs

STM32 Connectivity

STM32 Ecosystem



STM32 Performance Continuum

Microprocessors

Dual Cortex-A7 @ 650MHz *(2470 DMIPS)*
Cortex-M4 @ 200MHz *(250 DMIPS)*

Microcontrollers

Cortex-M7 @ 400MHz *(856 DMIPS)*

Cortex-M7 @ 216MHz *(462 DMIPS)*

Cortex-M4 @ 180MHz *(225 DMIPS)*

Cortex-M3 @ 120MHz *(150 DMIPS)*

STM32 High Performance Portfolio

STM32MP1 Hardware

life.augmented

STM32MP1 Block Diagram

STM32MP157

System
- 5x LDOs
- Crystal & Internal oscillators
- MDMA + 2x DMA
- Reset and Clock
- Watchdogs (2x I & W)
- 96-bit unique ID
- Up to 176 GPIOs

Security
- TrustZone
- DES, TDES, AES-256
- SHA-256, MD5, HMAC
- 3x Tamper Pins with 1 active
- T°, V and 32KHz detection
- Secure ROM and RAMs
- Secure Peripherals
- Secure RTC
- Analog true RNG

Dual Cortex-A7 @ 650MHz
- Core 1 @ 650MHz L1 32KB I / 32KB D
- Core 2 @ 650MHz L1 32KB I / 32KB D
- NEON SIMD
- NEON SIMD
- 256KB L2 cache

Cortex-M4 @ 200MHz
- FPU
- MPU

DDR3/DDR3L 32-bit @ 533MHz
LPDDR2/LPDDR3 32-bit @ 533MHz

- System RAM 256KB
- MCU System RAM 384KB
- Retention RAM 64KB
- Backup RAM 4KB
- Boot ROM 128KB
- OTP Fuse 3Kb

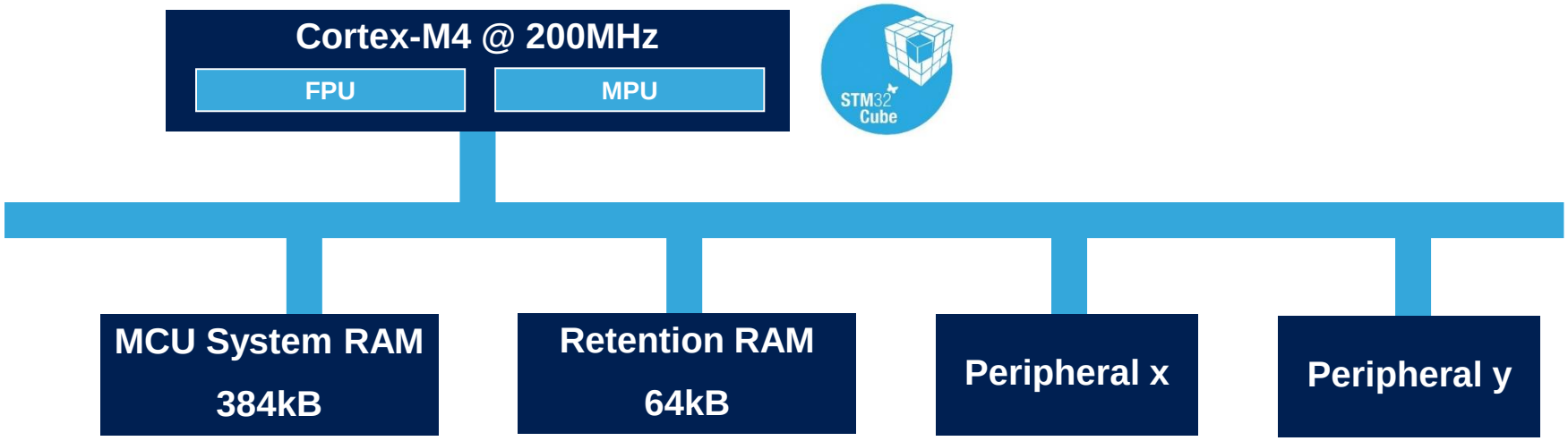
Control
- 2x 16-bit motor control PWM synchronized AC timer
- 10x 16-bit timer
- 5x 16-bit LP timer
- 2x 32-bit timer

3D GPU OpenGL ES2.0 @ 533MHz
26Mtri/sec, 133Mpix/sec

Connectivity
- 24-bit Parallel RGB Display
- MIPI DSI 2 lanes @ 1Gbps
- Camera Interface
- HDMI CEC
- 1Gbps Ethernet
- 2x FDCAN / TTCAN
- 2x USB2.0 Host HS
- USB2.0 OTG FS/HS
- MDIO
- DFSDM 8 channels / 6 filters
- 6x I²C
- 4x UART, 4x USART
- 6x SPI / 3x I²S
- 4x SAI
- SPDIF Tx / Rx 4 inputs
- Dual Quad SPI
- 3x SDIO3.0 / SD3 / eMMC 4.51
- 16-bit SLC NAND, 8-bit-ECC

Analog
- 2x 16-bit ADC 3MSPS 22 channels
- 2x 12-bit DAC
- Temperature sensor

Cortex-M4 MCU Sub System



➔ STM32 MCU with dedicated 448kB SRAM inside STM32MP1

➔ STM32 Cube Firmware full reuse on Cortex-M4 MCU Sub System

Graphics

- 3D GPU
  - OpenGL ES 2.0 support
  - 533MHz
  - 26Mtri/sec, 133Mpix/sec
  - Native support for OS like Linux, Android, Android Wear
  - HMI UI composition with 3D effects
  - Image / Video processing capable (3D effects, filtering, augmented reality, …)
- Display controller
  - MIPI DSI Tx 2 lanes @ 1Gbps
  - 24-bit RGB888
    - WXGA (1366x768) @ 60fps resolution support
    - 2 display layers with programmable colour LUT
    - Support for RAM less display

STM32MP15x Part Numbers

ARM® Cortex®-A7 – 650 MHz

ACCELERATION
- Dual core* ARM® Cortex®-A7 processor
- L1 and L2 caches
- 3D Graphic Processing Unit*
- Floating Point Unit + ARM® NEON™
- ARM® Cortex®-M4 200MHz coprocessor
- MDMA + DMA
- LPDDR2/LPDDR3 16/32**-bits 533MHz
- DDR3/DDR3L 16/32**-bits 533MHz

CONNECTIVITY
- 2 x USB2.0 HS Host
- USB2.0 OTG FS/HS
- 3 x SDMMC/SDIO
- USART, UART, SPI, I²C
- 2 x (TT)FD-CAN2.0*
- Gigabit Ethernet IEEE 1588***
- FMC (NAND Flash)
- Camera I/F
- Dual mode Quad-SPI
- DSI 2Gbit/s*

Product line	Cortex®-A7 cores	f_{CPU} (MHz)	Cortex®-M4 core	f_{MCU} (MHz)	f_{GPU} (MHz)	HW Crypto	3D GPU	FDCAN	MIPI®-DSI
STM32MP151A	1	650	1	200	-				
STM32MP151C						●			
STM32MP153A	2	650	1	200	-			2	
STM32MP153C						●			
STM32MP157A	2	650	1	200	533		●	2	●
STM32MP157C						●			

Notes: * not available in all product lines
** 16/32-bits for LFBGA448 and TFBGA361 packages, 16-bits only for LFBGA354 and TFBGA257 packages
*** 10/100M Ethernet only for LFBGA354 and TFBGA257 packages

life.augmented

STM32MP1 PMIC companion IC: STPMIC1

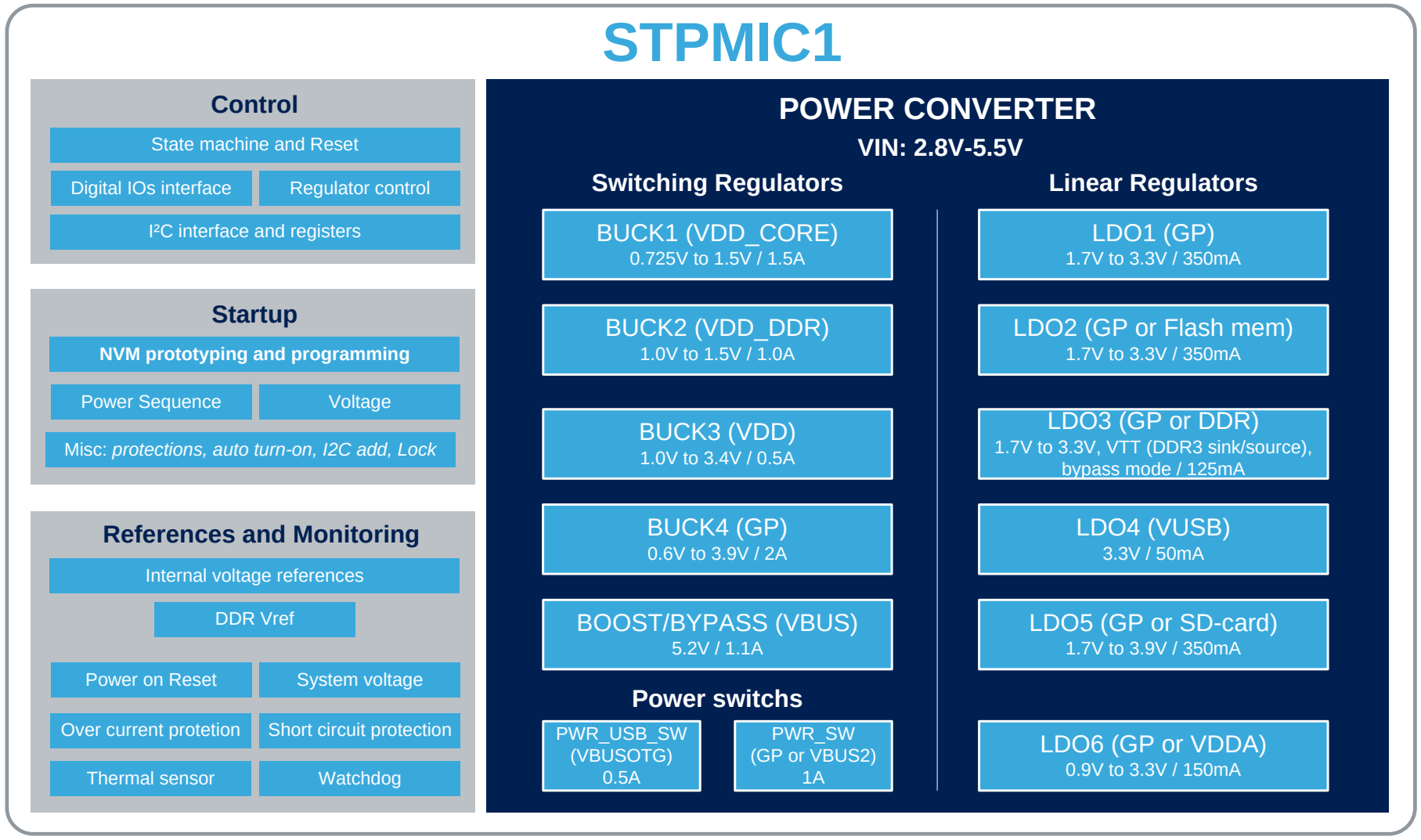
Why STPMIC1?

Optimized Power Consumption

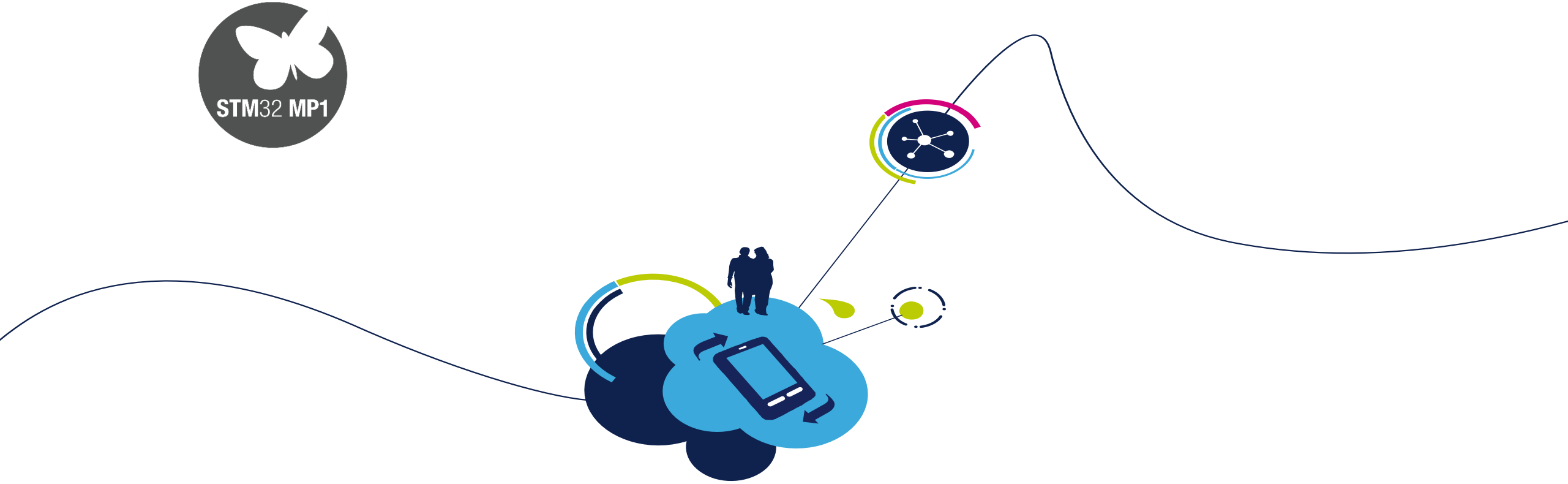
BOM savings on Typical Applications

Low PCB footprint vs. full discrete solution

STPMIC1 Block Diagram



Package: WFQFN 5x6x0.8 44L pitch 0.4mm



STM32MP1 Embedded Software

life.augmented

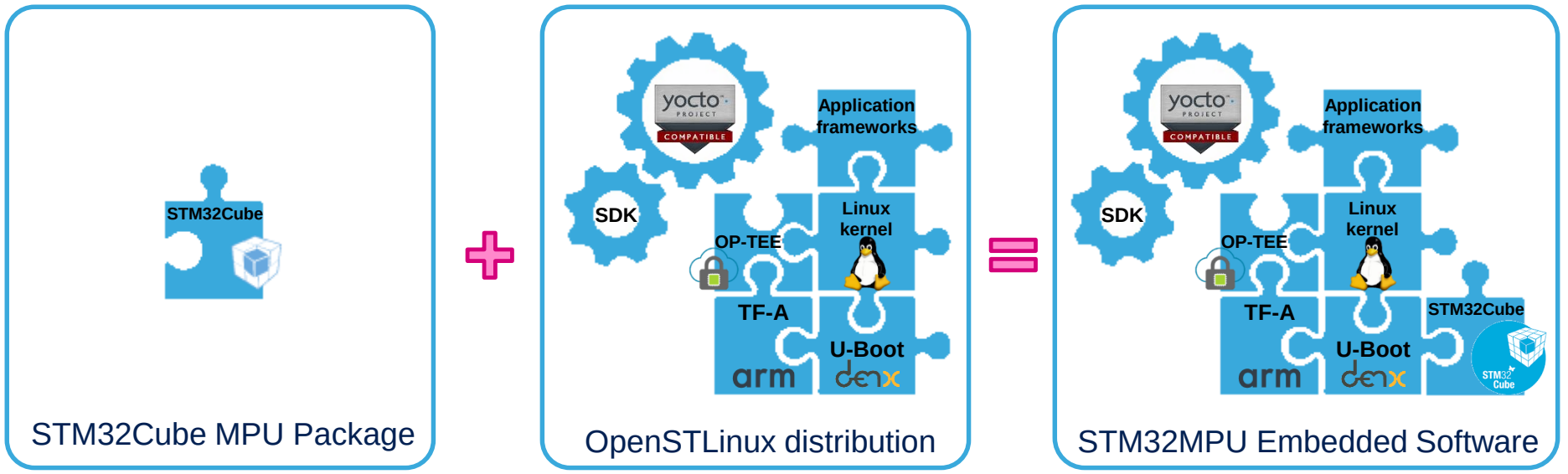
Linux & Real Time OS

- Full development environment for Linux and RTOS
  - Linux running on Dual Cortex-A7 cores: Bootloader, Kernel, User Space
  - Cube running on Cortex-M4 core: MCU co-processor firmware
- Build and debug environments for different user profiles:
  - Starter Package
    - to quickly and easily start with any STM32MP microprocessor device
  - Developer Package
    - to add your own developments on top of the STM32MPU Embedded Software distribution
  - Distribution Package
    - to create your own Linux® distribution, your own Starter kit and your own Developer kit
- Drivers for all peripherals available on Linux Software and/or MCU co-processor firmware sides

STM32MPU Embedded Software

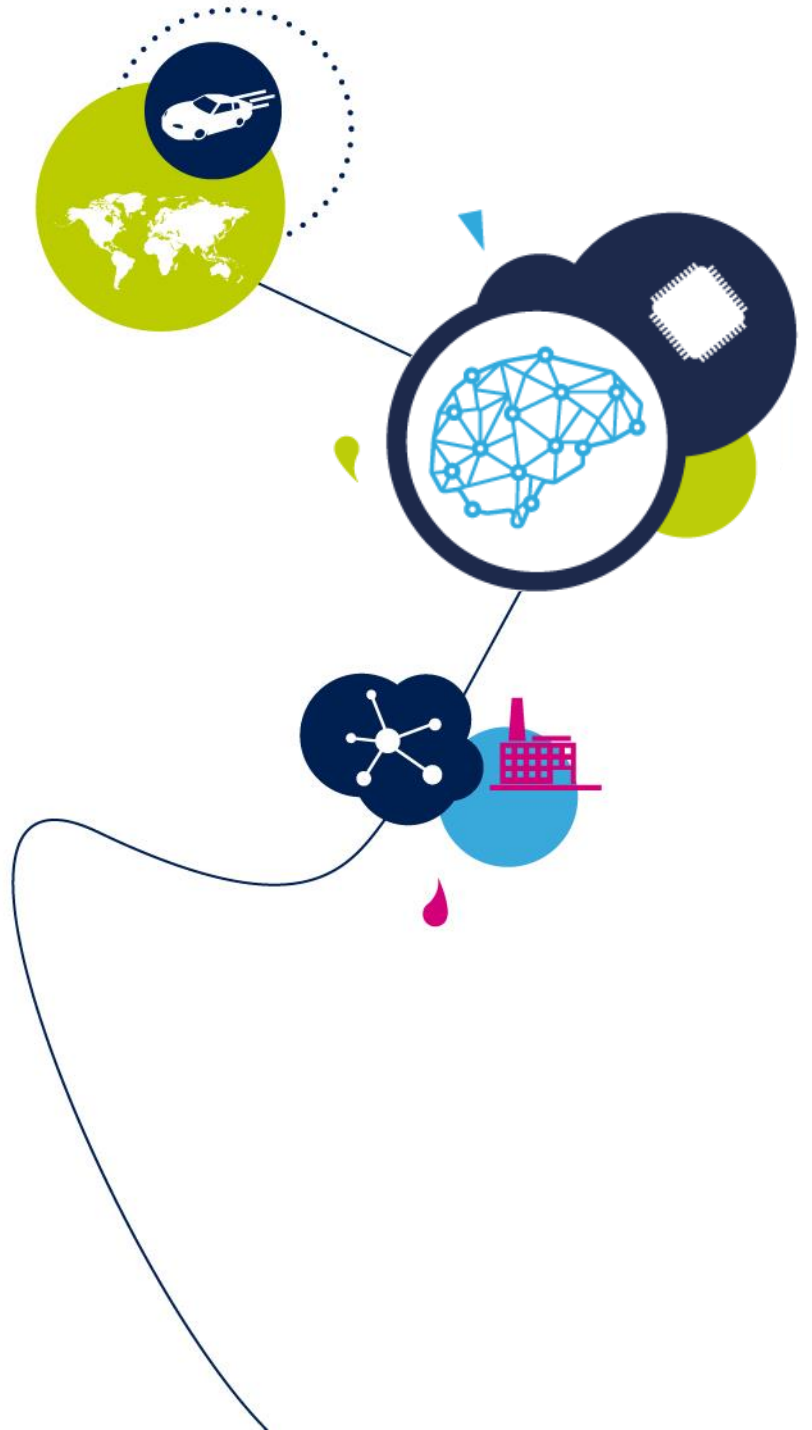
Cortex-M4 core

Dual Cortex-A7 cores



OpenSTLinux

ST MEMS & SENSOR



life.augmented

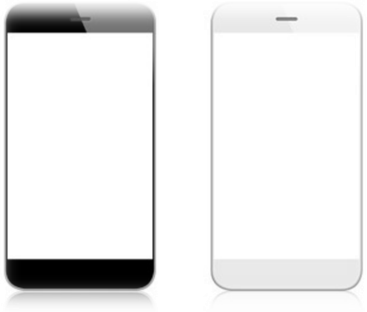
15B
Shipped Sensors

6Mu/d
Best in class service

ST sensor in your life

Personal Electronics

1 ST MEMS every 2 phones



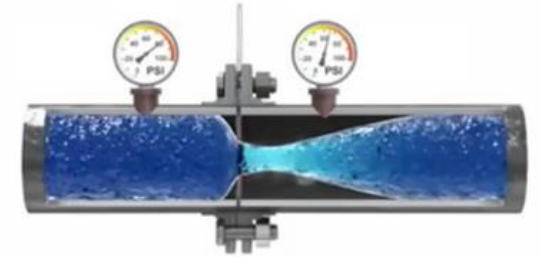
Automotive

1 ST MEMS every 3 car navigators

life.augmented

Industry

Leadership in WP and Harsh environment pressure sensor



Computing

1 ST MEMS every 2 DRAM

SENSORS & MOTION MEMS

ST offer

	Available in 2018	Applications
CONSUMER	AXL, 6-axis IMU, Mag, E-compass, Microphone, Pressure, Humidity, Temperature	IOT, Wearable, Alarm, Smart Home, Remote Control, Voice Assistant
INDUSTRIAL	AXL, 6-axis IMU, Mag, E-compass, Microphone, Dedicated AXL	Indus Robot, Positioning, Tracking, Tilt, Vibration
AUTOMOTIVE	AXL, Gyro, 6-axis IMU	Alarm, E-call, Telematic, Vehicle tracking

- CONSUMER: Features; Wide offer, intro of new products; Competitive price
- INDUSTRIAL: High Accuracy; Dedicated products; 10 Years Longevity
- AUTOMOTIVE: AEC-Q100; >10 Years Longevity

We create the new generation of Sensors to allow developers to exploit their potential by improving the overall system efficiency thanks to:

- Reduced Power Consumption
- Increased Performance

And we do this by leveraging Machine Learning techniques from the world of A.I.

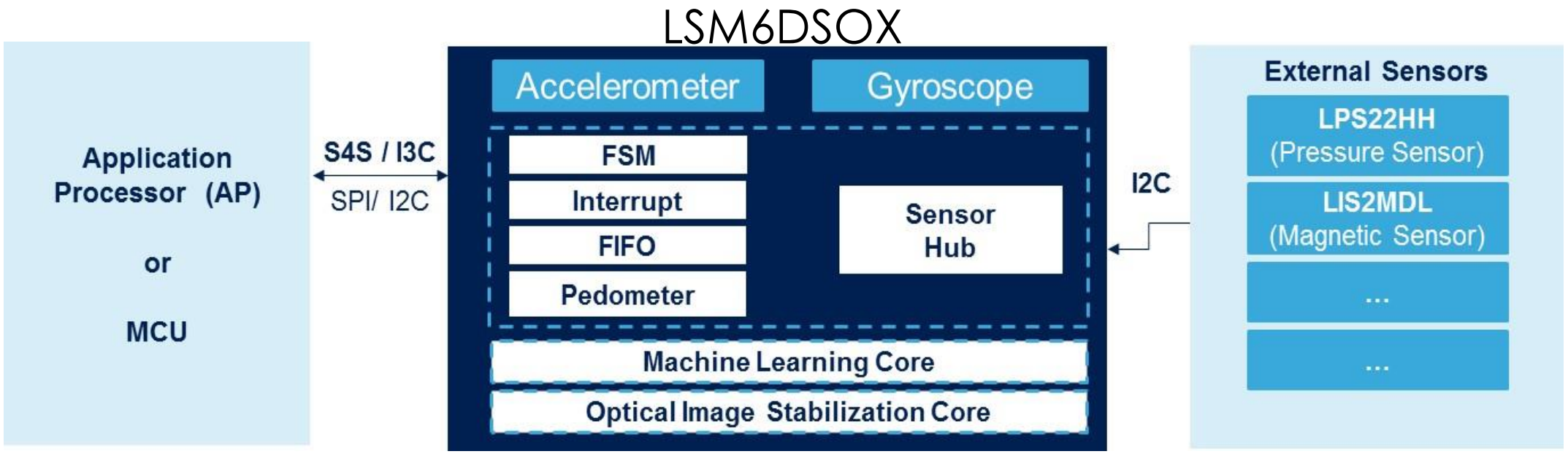
from Low Power Sensor to Low Power System

Best in class IMU **Power Consumption** (0.55 mA combo mode HP)

10 to 1000 time energy saving by running Machine Learning on Sensor HW (vs. AP)

Simplified And Lean System Concept based on 2 mainstreams:

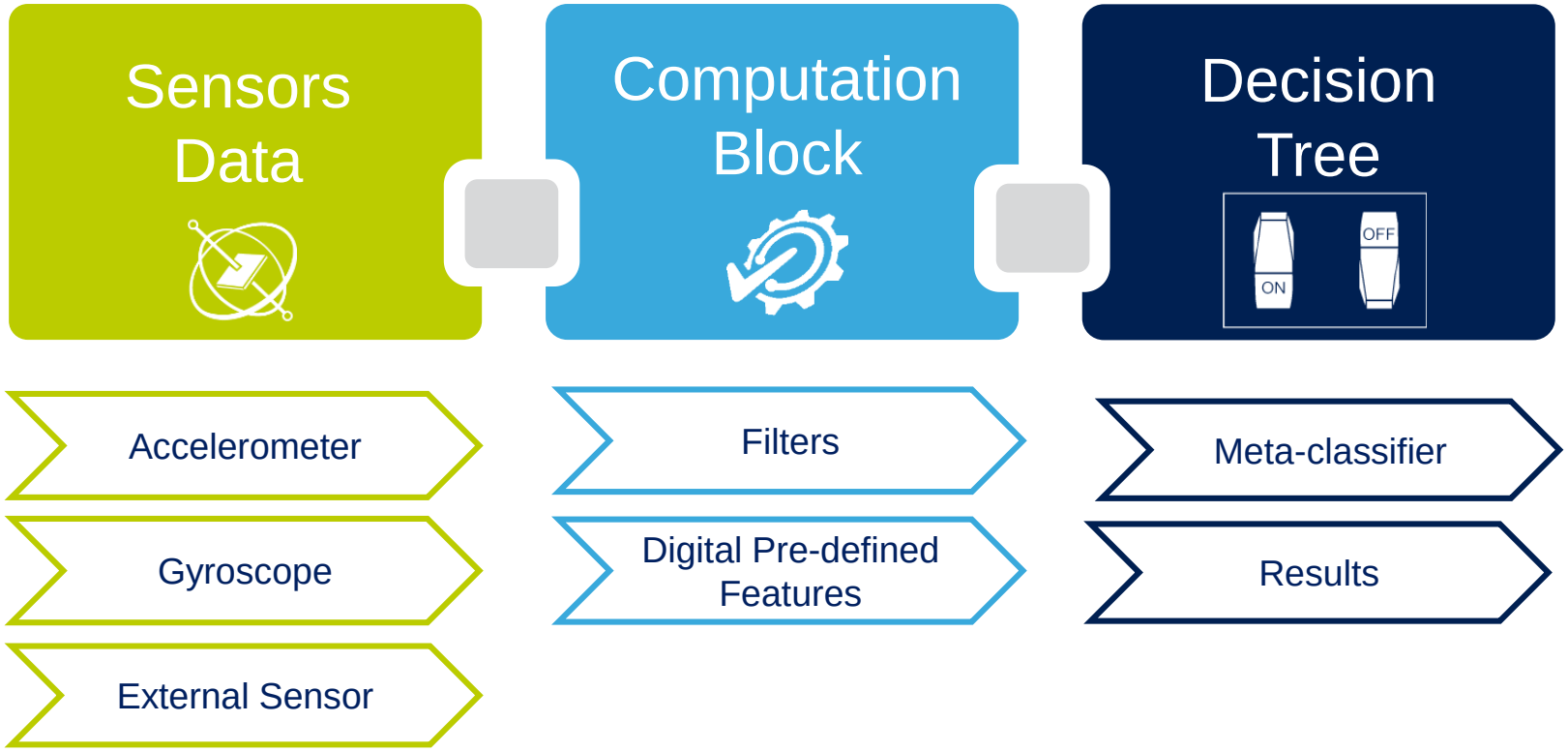
1. Configurable power mode and high speed communication
2. Flexible HW solution running accurate algorithms



Note: FSM→ Finite State Machine

Machine Learning Core (MLC)

10 to 1000 time energy saving by running Machine Learning on Sensor HW (vs. AP)



life.augmented

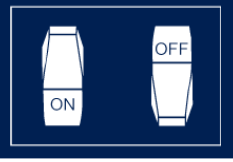
Machine Learning Core (MLC)

10 to 1000 time energy saving by running Machine Learning on Sensor HW (vs. AP)

Sensors Data

Computation Block

Decision Tree



A wide Set of Inputs selectable

1. Accelerometer $[a_x\ a_y\ a_z]$, $[a_v]$, $[a^2_v]$
2. Gyroscope $[g_x\ g_y\ g_z]$, $[g_v]$, $[g^2_v]$
3. External sensor $[m_x\ m_y\ m_z]$, $[m_v]$, $[m^2_v]$

Magnitude Available

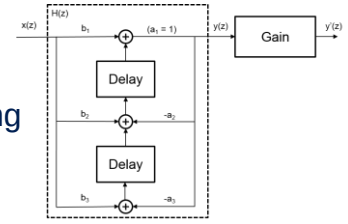
$$V = \sqrt{X^2 + Y^2 + Z^2}$$

Filter Block and Computation

Advanced computation features

Mean, Variance, Energy, Peak, etc

Data filtering



Activity recognition … and More

life.augmented

Machine Learning Core Programming

5 simple steps to **improve up to 20% the detectability** by using A.I. based solutions

1

User defines Classes to be recognized

2

Collect Logs for each class

3

Define Features that best characterize the identified classes

4

Machine Learning tools generate program for LSM6DSOX based on Logs and Features

5

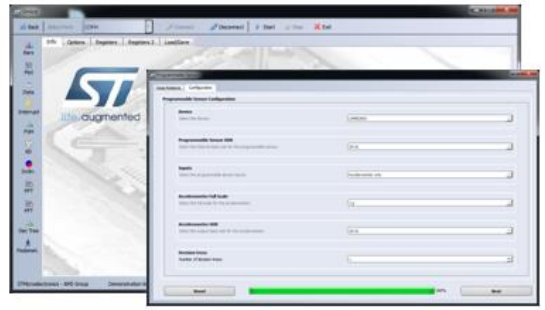
Configure the LSM6DSOX and run the application

This is the added Value!

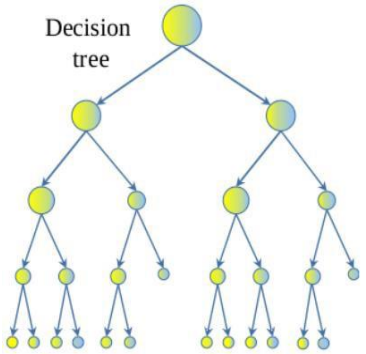
life.augmented

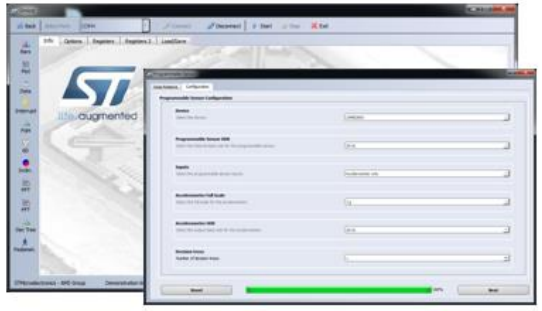
Machine Learning Core - PROCESS

1. Define classes to be recognized
   - (e.g. Activity Recognition: Walking, Jogging, Driving)
2. Collect multiple logs for each class
   - (e.g. different persons doing the same activity)
3. Offline data analysis:
   - Define the features that better characterize the defined classes (some statistical parameters like: mean, variance, energy, Peak to Peak)
   - Implement **decision tree** (threshold, number of nodes, outputs, etc…) using machine learning tools WEKA
4. Configure LSM6DSRX / LSM6DSOX
   - **Decision tree** will run on the device, reducing MCU Power Consumption



Unico





Unico

life.augmented

End customer use case examples

The Machine Learning Core identifies if a data pattern (motion, pressure, temperature, magnetometer, etc) matches an activity in a user defined set of classes, for instance it recognize if you are Running, Walking, Driving, in an Airplane, etc

Activity recognition

Stationary, walking, fast walking, jogging, biking, driving.

Gym activity recognition

Bicep curls

Squats

Count the number of bicep curls, squats and push-ups, etc

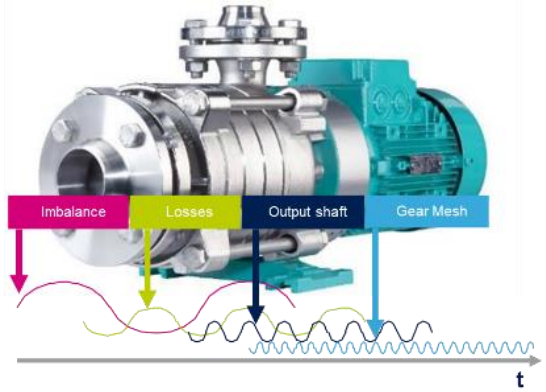
Airplane Mode detection

Take off

Interrupt generation

Recognize Take-off and Landing to set the Smartphone (Radio off)

Predictive Maintenance



Data collection, processing, Analytics with FFT, Vibration, Sensors fusion data

life.augmented

FSM / MLC

Power Consumption comparison

FSM and MLC are incredibly efficient in current consumption needs:

+ ~3 µA for each Finite State Machine

+ 1-15 µA Machine Learning Core

MLC Algorithm example	Sensors used	ODR	Number of decision trees	Number of nodes	MLC add. Current consumption
Vibration Monitoring	A	26 Hz	1	2	1 µA
Motion Intensity	A	12.5 Hz	1	7	1 µA
6D position recognition	A	26 Hz	1	8	2 µA
Activity Recognition for mobile	A	26 Hz	1	126	4 µA

Trilobyte: SensorTile.Box*

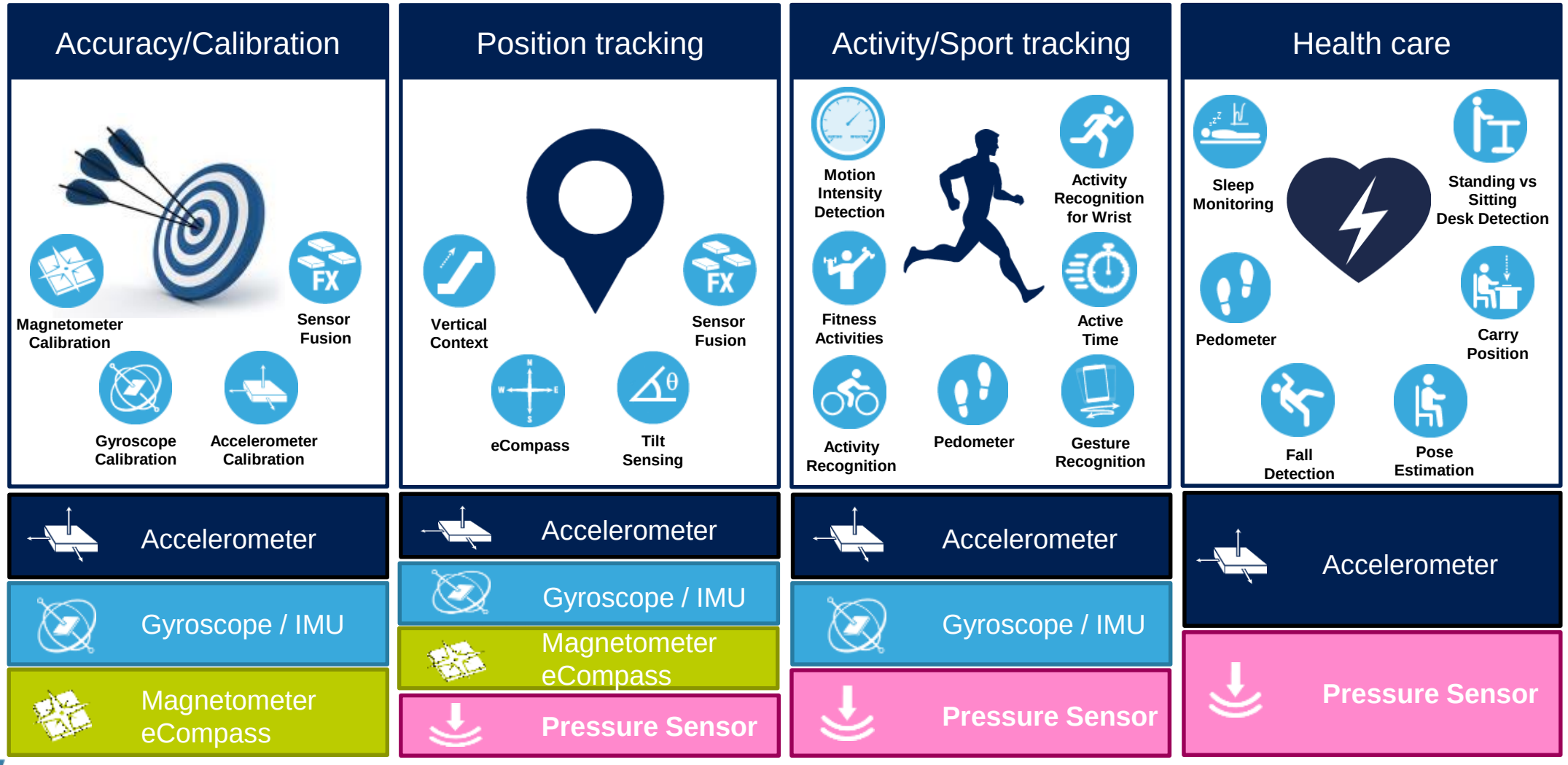
- SensorTile.Box is a Sensors unit
  - A ready-to-use measurement unit in Plastic Case (57x38x20) IP54
  - A mobile phone App
- Sensors:
  - STTS22H Accurate temperature (once available)
  - LSM6DSOX (Low power accurate 6x IMU with MLC)
  - IIS3DHH (Inclinometer)
  - IIS2DLPC (low power, high performance Accelerometer)
  - IIS2MDC (Magnetometer)
  - LPS22HH (Altimeter / pressure sensor)
  - MP23ABS1 (Analog wide-band microphone)
  - HTS221 (Humidity sensor, HTS2 once available)

- The Sensor Kit has 3 operational modes
  - Beginner -> provides predefined functions
  - Developer -> Configure and define new Apps
  - Expert -> Develop specific Solutions with a coding approach
- Other ST components
  - STM32L4R9
  - STBC02 (Battery charger)
  - SPBTLE-1S (BLE Module - to be replaced with Module 2 once ready)

life.augmented

*Available during Q2

MotionSW available for your Sensor



All libraries are available in X-CUBE-MEMS1, X-CUBE-MEMS-XT1 package

life.augmented

SENSORS with MLC

- LSM6DSOX
  - 6-axis IMU for IOT & Wearable
  - 8x MLC working in parallel
- LSM6DSRX*
  - 6-axis IMU for Telematics & Robotics
  - 8x MLC working in parallel
- ISM330DHCX*
  - 6-axis IMU for Industry 4.0
  - 8x MLC working in parallel
- IIS2ICLX*
  - 3-axis Inclinometer for infrastructure monitoring
  - 8x MLC working in parallel



IOT / Wearable
Movement tracking & Shock

GNSS, Telematics, VR/AR, Robot
Rotation / Movement


Predictive maintenance & Monitoring
Vibration / Tilt




infrastructure & Landslide monitoring
Mechanical movement / Tilt

* Available during 2019

life.augmented


life.augmented