Unboxing the MG24 and AI/ML Foundations

May 2022 | Antonio Trujillo Rojas

Agenda

1. xG24 Overview and xG24-DK2601B dev kit
2. Low-power peripherals and MVP (AI/ML accelerator)
3. Introductory concepts to AI/ML
4. Simplicity Studio v5
5. Lab session

An Overview of the EFR32xG24 & Dev Kit

High Performance, Low-Power and Secure 2.4 GHz Wireless SoC

BG24 and MG24: Optimized for Battery Powered IoT Mesh Devices

Sensing at the Edge

AI/ML Hardware Accelerator Key Features

- Optimized Matrix processor to accelerate ML inferencing with a lot of processing power **offloading the CPU**
- Real and complex data
- **2x to 4x faster** inferencing over Cortex-M
- Up to **6x lower power** for inferencing

Low-Power SoCs and Modules Optimized for Battery Powered IoT Mesh Devices

High Performance Radio
-Up to +19.5 dBm TX
-97.6 dBm RX @ BLE 1 Mbps
-105.7 dBm RX @ BLE 125 kbps
-104.5 dBm RX @ 15.4
Improved Wi-Fi Coexistence
RX Antenna Diversity

Low Power
5.0 mA TX @ 0 dBm
19.1 mA TX @ +10 dBm
4.4 mA RX (BLE 1 Mbps)
5.1 mA RX (15.4)
33.4 µA/MHz
1.3 µA EM2 with 16 kB RAM

World Class Software
Simplicity Studio 5
Matter[1]
Thread[1]
Zigbee[1]
Bluetooth (1M/2M/LR)
Bluetooth mesh
Dynamic multiprotocol[1]
Proprietary

ARM® Cortex®-M33
78 MHz (FPU and DSP)
Trustzone®
Up to 1536kB of Flash
Up to 256kB of RAM

Dedicated Security Core
Secure Vault™ - Mid
Secure Vault™ - High

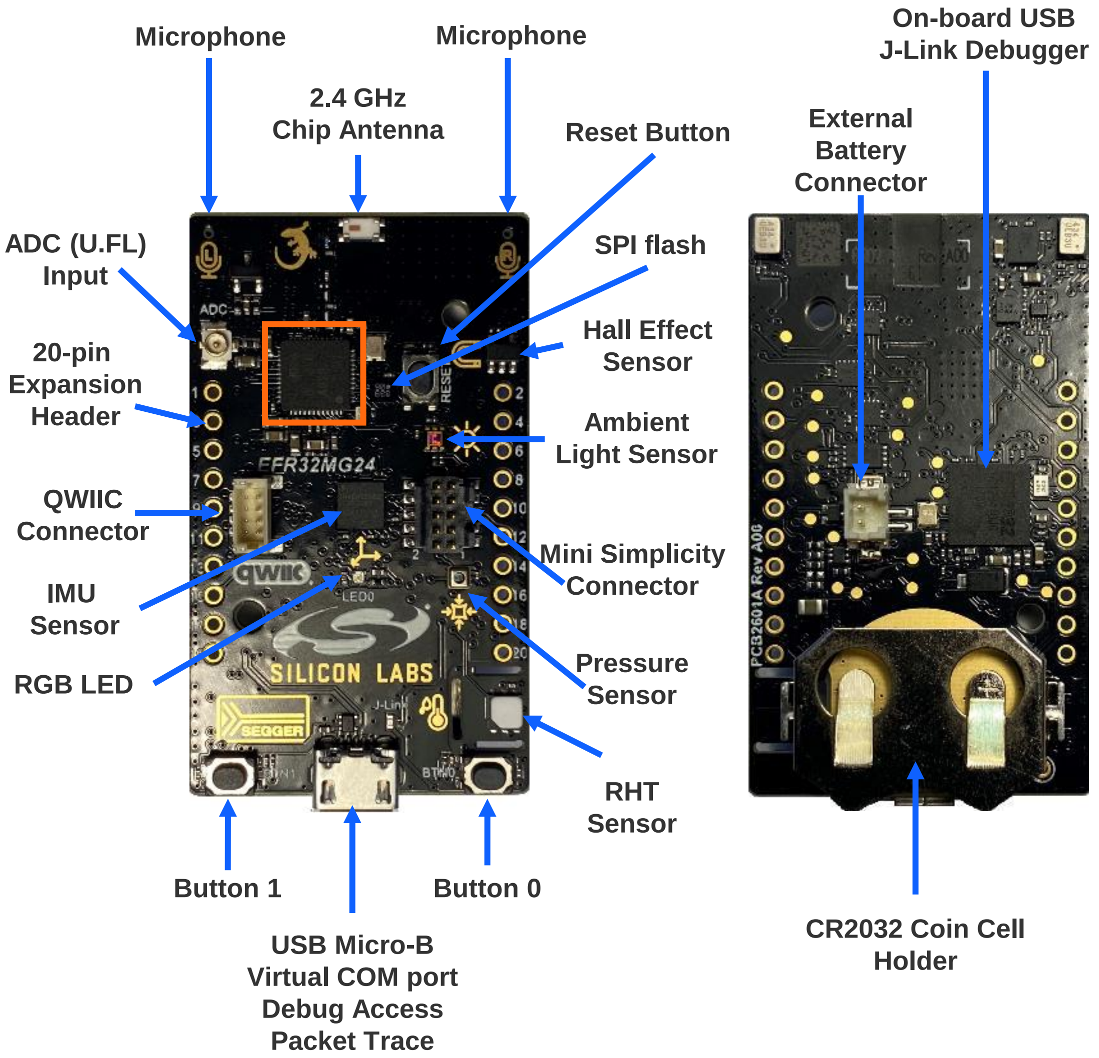
Low-power Peripherals
EUSART, USART, I2C
20-bit ADC, 12-bit VDAC, ACMP
Temperature sensor +/- 1.5°C
32kHz, 500ppm LFRCO

AI/ML
AI/ML Hardware Accelerator

SoCs and Modules
5x5 QFN40 (26 GPIO) -125ºC
6x6 QFN48 (28/32 GPIO) -125ºC
7x7 SiP Module (+10 dBm)
12.9x15.0 PCB Module (+10 dBm)

[1] Requires MG24

xG24-DK2601B Dev Kit: a Powerful Prototyping Platform



Features

- EFR32MG24B310F1536IM48
  - 1536 kB Flash and 256 kB RAM
- Wireless SoC with multi-protocol radio

Advanced Features

- AI/ML Hardware Accelerator - MVP
- 20-bit ADC

Broad Range of Sensors

- I2C
  - RHT Sensor (*Si7021*)
  - Hall-effect Sensor (*Si7210*)
  - Pressure Sensor (*BMP384*)
  - Ambient Light Sensor (*VEML6035*)
- SPI
  - 6-axis IMU (*ICM-20689*)
- I2S
  - 2x MEMS Digital Microphones (*ICS-43434*)

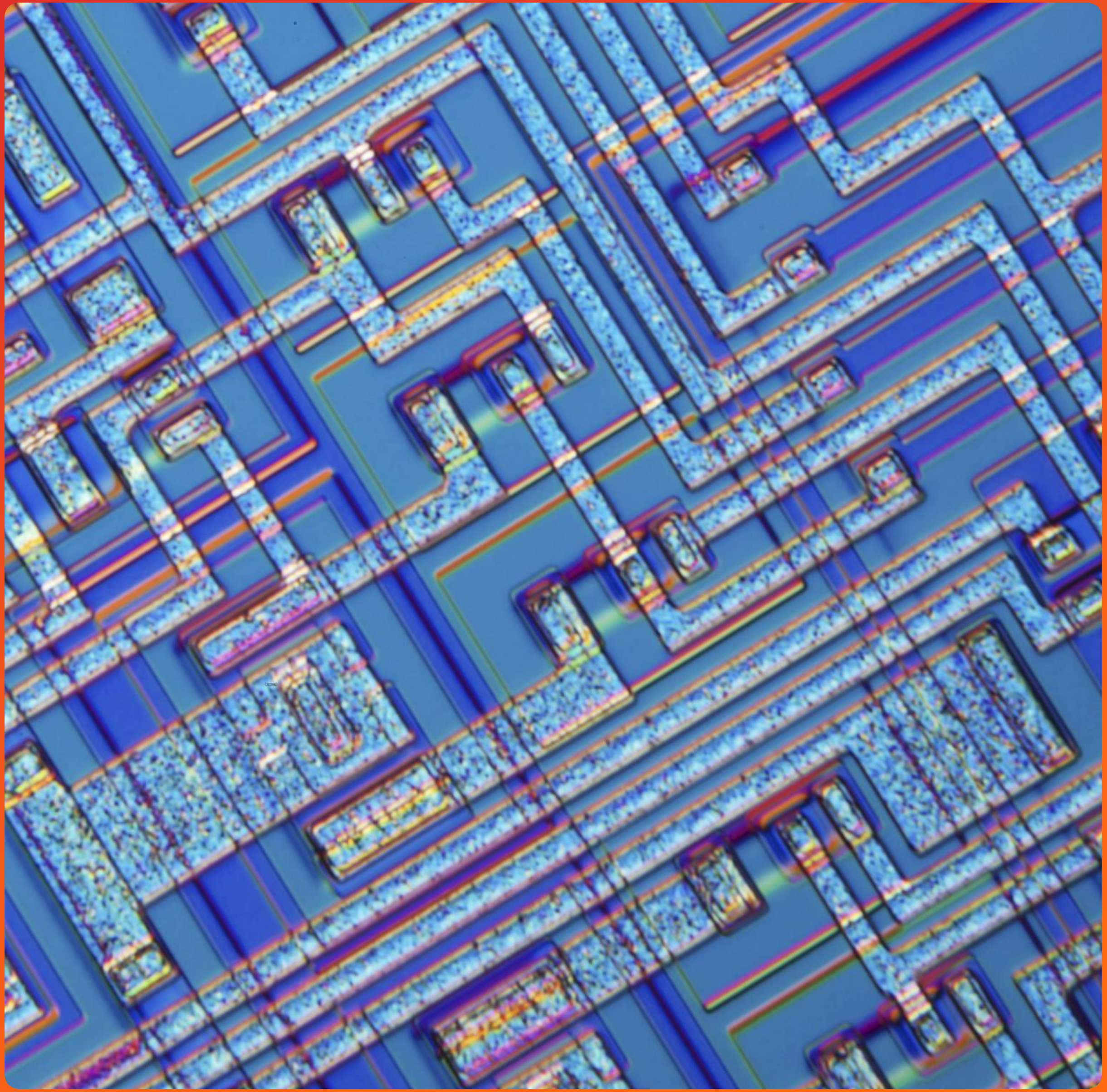
Connectors:

- Breakout pads
- QWIIC connector
- External battery connector
- Coin Cell battery Holder
- Mini-Simplicity connector
- Micro USB (power, vcom, debug, PTI)

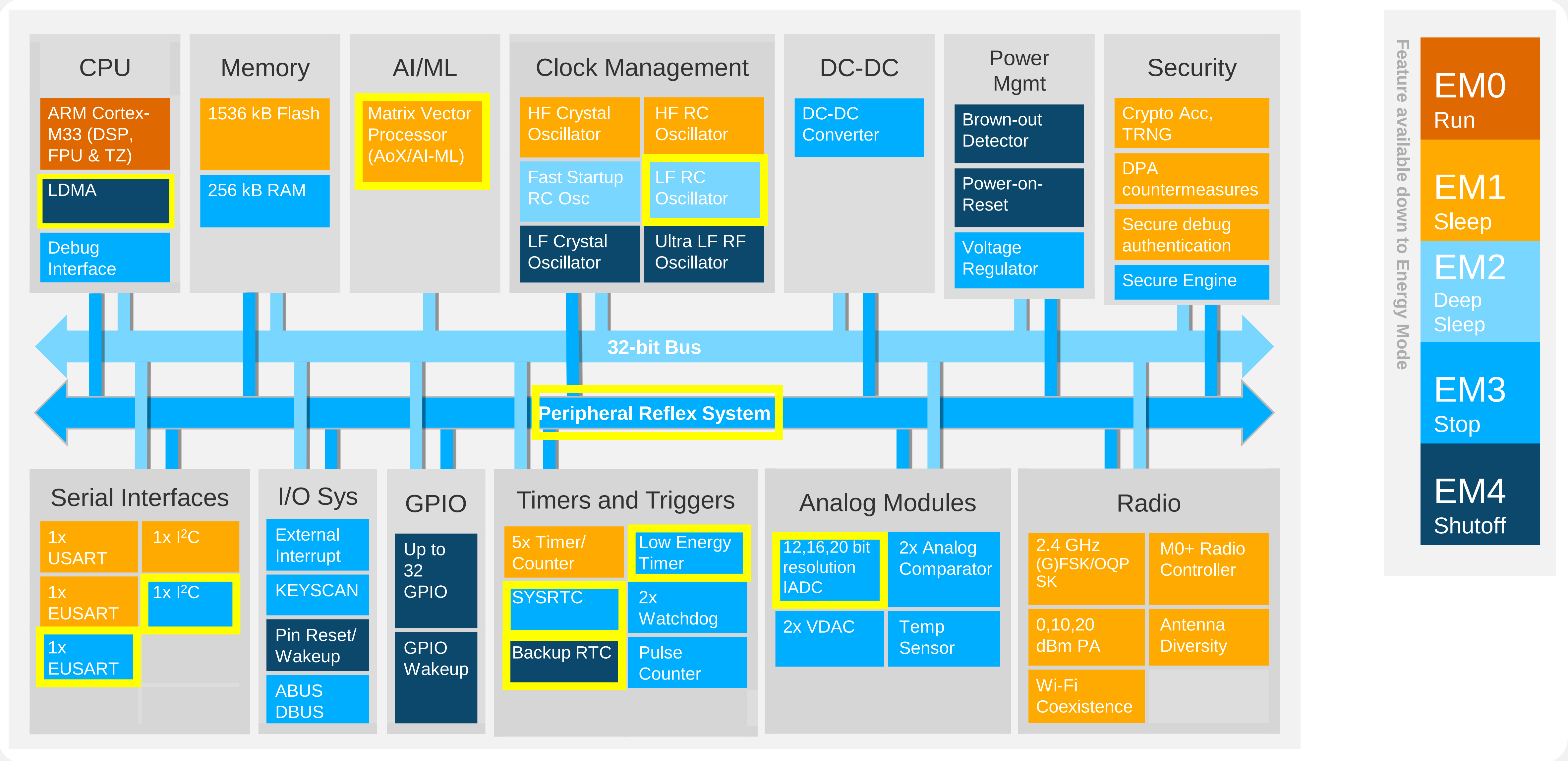
User Interface

- RGB LED and Push Buttons
- SPI flash
- U.FL connector for dedicated iADC input
- Precise External VREF for iADC – 1.25v ± 0.12% (ADR1581)

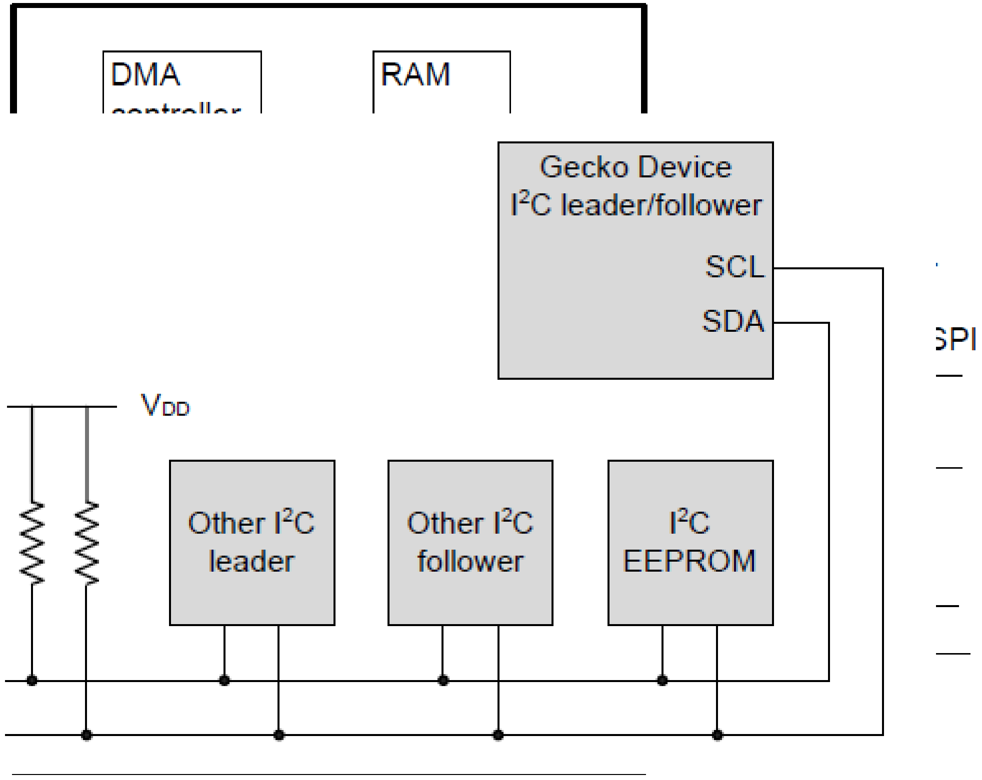
xG24 Low-Power Peripherals and MVP (AI/ML Accelerator)



xG24 Block Diagram

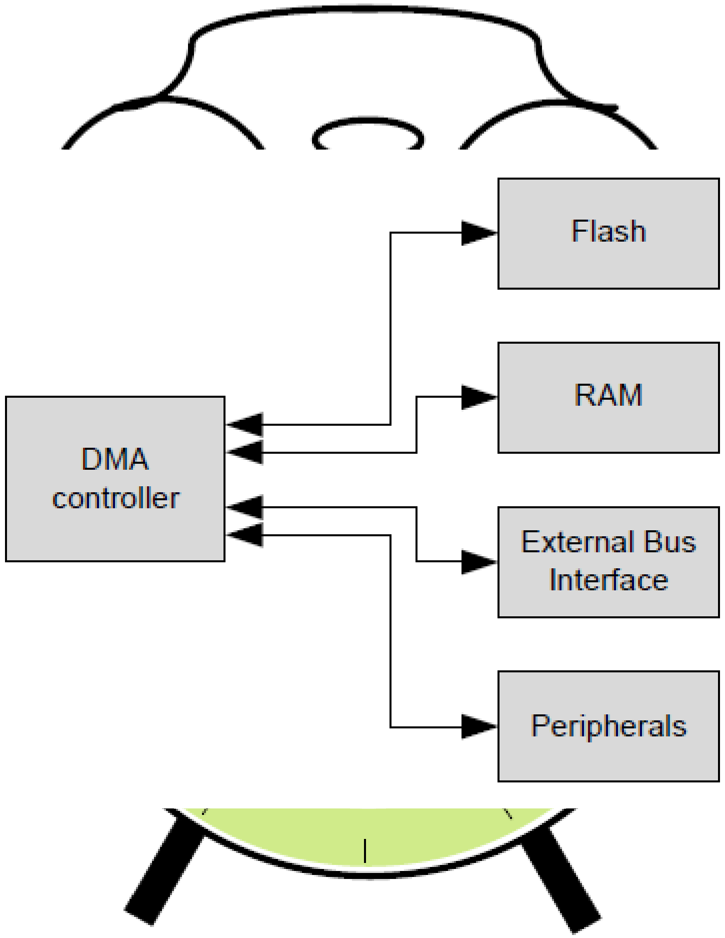


xG24 Low-Power Peripherals



- **LFRCO**
  - Integrated fast start-up RC oscillator
  - Precision mode
    - ± 500 ppm accuracy (self calibrating vs HFXO)
    - Available in EM2
    - Target: BLE apps with no LFXO (reduce BOM cost)
      - Further details: docs.silabs.com
- **Serial communication**
  - EUSART
    - Asynchronous (UART, IrDA)
      - Wake from EM2 on special frame, RX timeout or watermark level in FIFO
      - LDMA transfer (EM1) on watermark level in FIFO
      - Max baud rate in low-energy mode 9600
    - Synchronous (SPI)
      - Secondary mode only
      - Wake or LDMA transfer (EM1) on watermark level in FIFO
      - Down to EM3
      - Maximum clock speed 10 MHz
    - Low energy mode available in EUSART0
  - I2C
    - Wake from EM3 on address recognition
    - LDMA transfers supported

xG24 Low-Power Peripherals - Continued

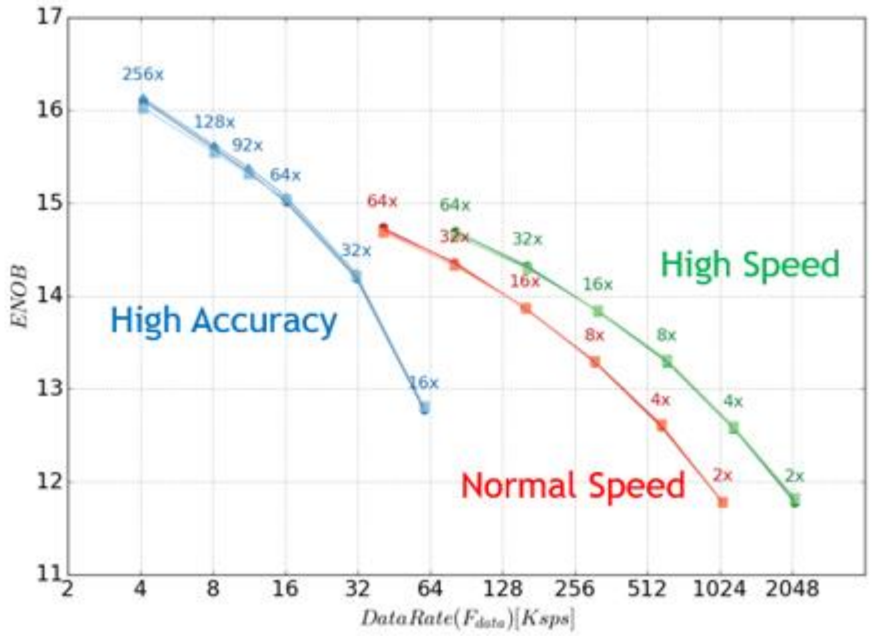


- **Low-Energy timers:**
  - **SYSRTC**
    - 32-bit length
    - Available in EM3
    - Shared between cores (Each core has separate CC channel groups)
    - Replaces RTCC and PRORTC
  - **BURTC**
    - 32-bit length
    - Available in EM4 (ULFRCO)
    - Single compare channel
  - **LETIMER**
    - 24-bit length
    - Available in EM3
    - 2 compare channels and PWM output
- **PRS**
  - Interconnect peripherals: “Producers” to “consumers”
  - Trigger peripheral operation autonomously
  - Logic operations on asynchronous channels (AND, OR, XOR…)
  - Available in EM3
- **LDMA**
  - Memory ↔ peripheral, Memory ↔ Memory, Peripheral ↔ Peripheral
  - Linked descriptors
  - Available in EM2/3 (Transfers occur in EM1)

xG24 Low-Power Peripherals - iADC 12-16 & 20 Bit Resolution

- **Normal Mode**
  - 12-bit output resolution, 11.7 ENOB @ 1 Msps (OSR = 2)
  - 16-bit output resolution, 14.3 ENOB @ 76.9 ksps (OSR = 32)
- **High-Speed Mode – Double speed, Similar Performance**
  - 12-bit output resolution, 11.7 ENOB @ 2 Msps (OSR =2)
  - 16-bit output resolution, 14.3 ENOB @ 153.8 ksps (OSR = 32)
- **High-Accuracy Mode – Highest performance**
  - Dedicated inputs for full performance across temperature
    - VREFN, VREFP, AIN0-3
  - 20-bit output resolution, 15 ENOB @ 15.3 ksps (OSR = 64)
  - 20-bit output resolution, 16 ENOB @ 3.8 ksps (OSR = 256)

Effective Number of Bits, External VREF



Oversampling Ratio (OSR) increases resolution and performance in all modes

MVP – Matrix Vector Processor (AI/ML Hardware Accelerator)

Benefits of processing AI/ML in device

- Lower power
- Save Bandwidth
- Lower Latency
- Ensure Privacy
- Higher Security
- Lower Cost

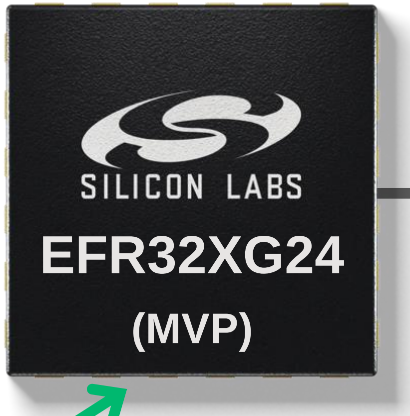
Use Cases for AI/ML

- Timeseries data on ADC or GPIO
- Audio mic array with beamforming
- Audio mic input with Audio Front End, DSP
- Image capture (incl. fingerprint reader)

Sensors
Acceleration, Temperature, Current/Voltage

Microphones
Analog or Digital

Camera
Low resolution imaging



Event Detection

AI/ML Hardware Accelerator Key Features

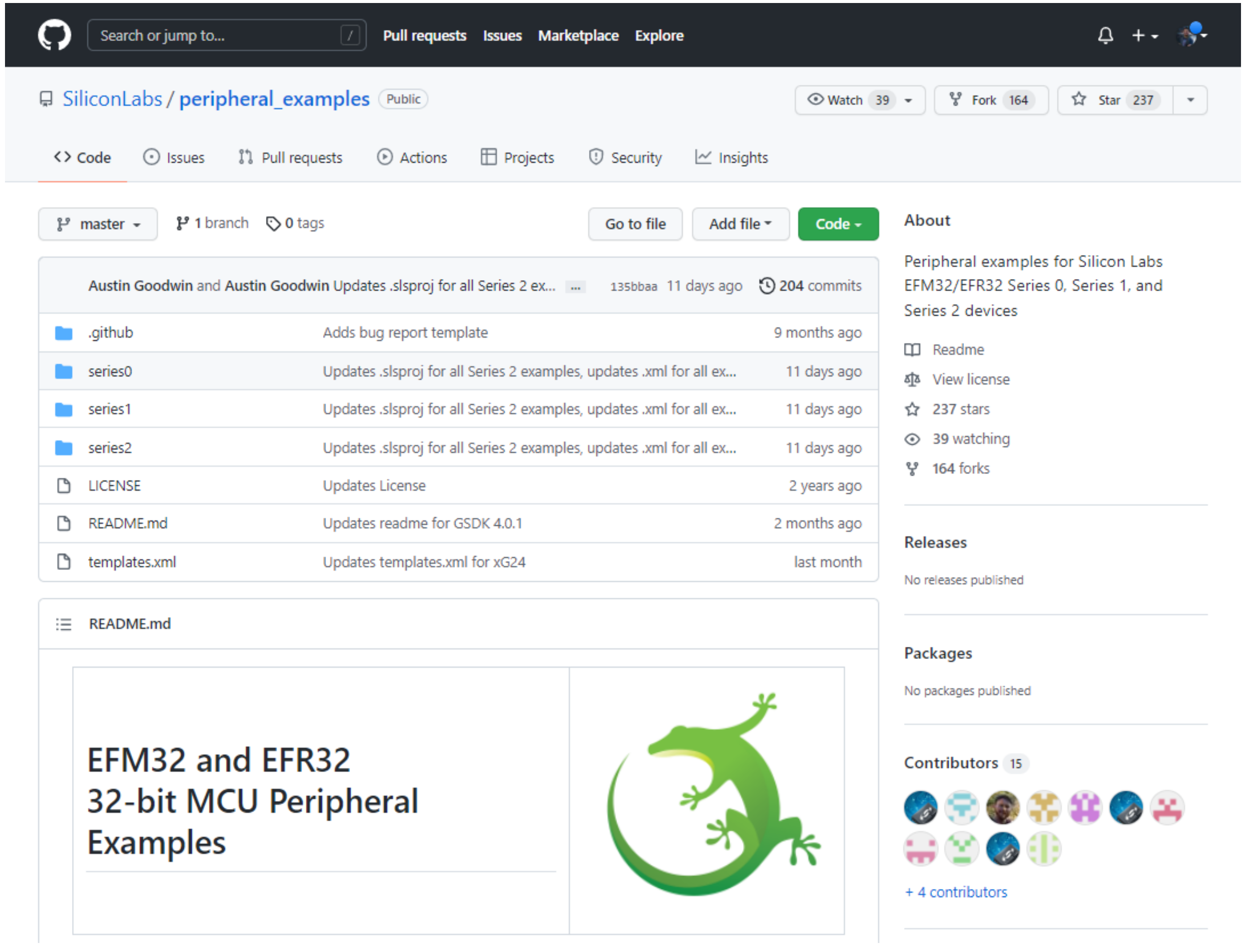
- Optimized Matrix processor to accelerate ML inferencing with a lot of processing power **offloading the CPU**
- Real and complex data
- **2x to 4x faster** inferencing over Cortex-M
- Up to **6x lower power** for inferencing

AI/ML Hardware Accelerator enables efficient Edge ML inferencing

Further Peripheral Information

- **EFR32xG24 reference manual**
  - Detailed information on peripheral operation
  - Details about peripheral registers
- **EFR32MG24 / BG24 datasheet**
  - Details regarding GPIO pinout
  - Details regarding peripherals available in different OPN
  - Details about peripheral functions available in each GPIO port
- **Peripheral examples GitHub repository**
  - Examples demonstrating functionalities of different peripherals
  - EMLIB based

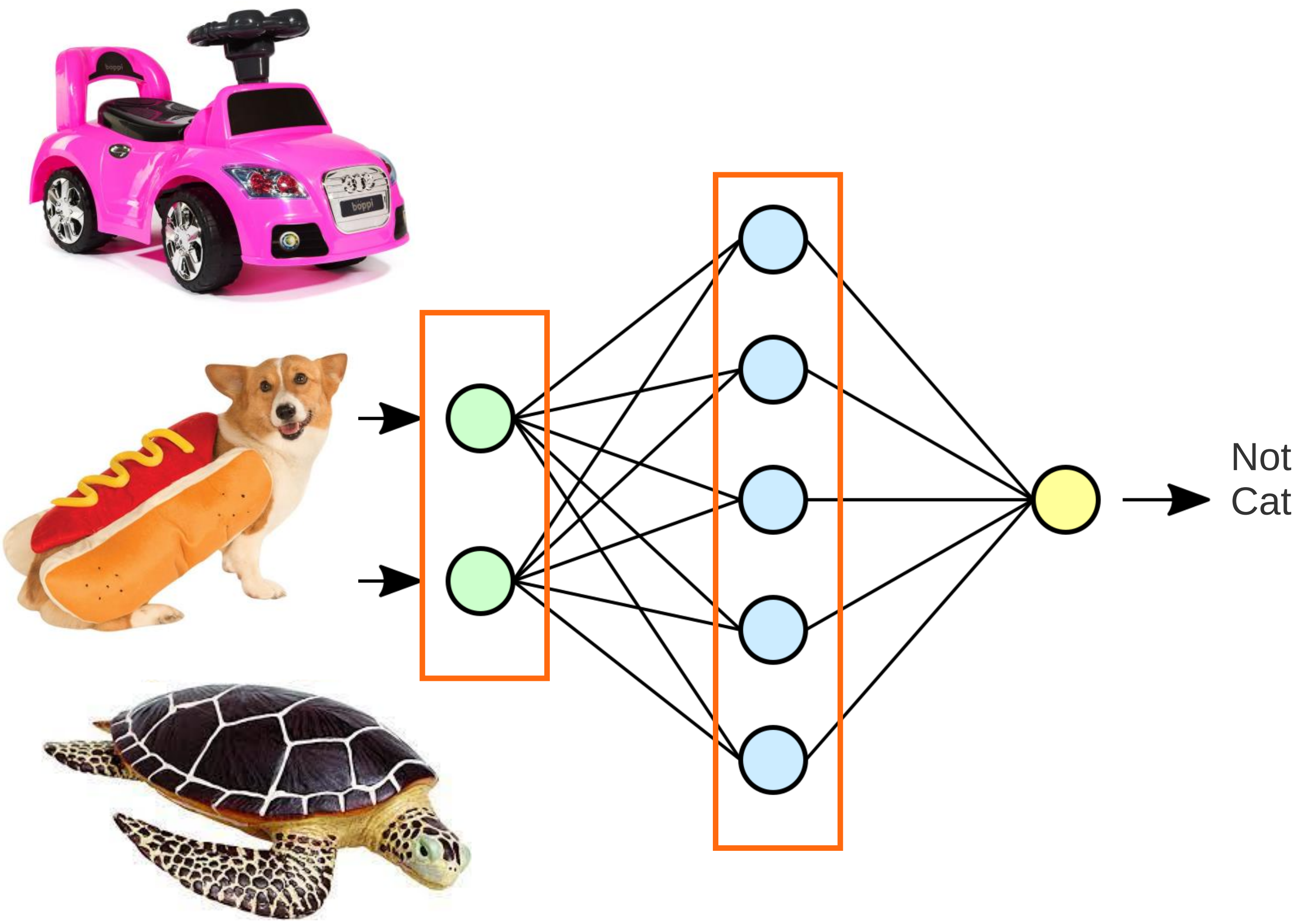
GitHub



Introductory Concepts to AI/ML

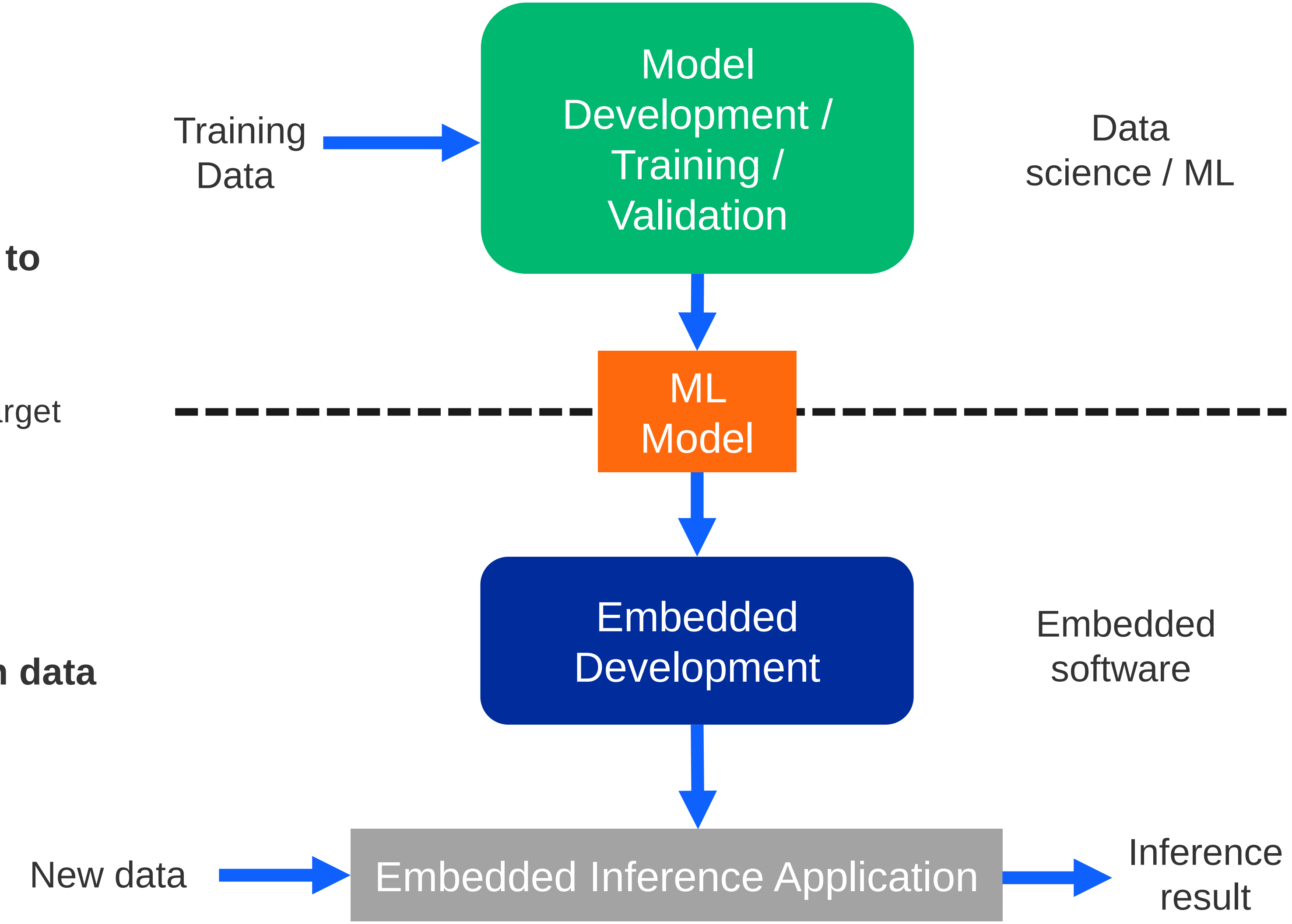
Machine Learning in a Nutshell

- **Classical Machine Learning**
  - Using training data in an automated process to adjust algorithms
  - Linear regression, support vector machines etc
- **Neural Networks**
  - Networks of basic operators mimicing the brain's neurons
  - Training adjusts operators so network gives expected result for a given input
- **Inference**
  - Using a trained model to process new data
- **Common usecases**
  - Supervised learning (e.g., Classification)
    - Labeled dataset
    - Classifying new input data to one of the classes the model was trained with
  - Unsupervised learning (e.g Anomaly detection)
    - No labelled data
    - Trained on "normal operation" dataset
    - Inference detects anomalies in input data that are very different from the "normal operation" dataset it is trained for



Neural Networks on Embedded Devices

- **Model trained on PC/server**
  - Very resource demanding
  - Training data set used as input
    - Preprocessed raw data for feature extraction
- **Trained model (flatbuffer) deployed to embedded target**
  - Used to run inference on new data
  - Normally no further training is done on target due to power/resource constraints
- **ML model is the connection point**
  - Data science / ML domain
  - Embedded software domain
- **Exact same pre-processing used on data in training must be applied in the embedded application**

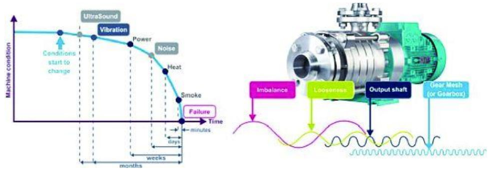


Artificial Intelligence(AI) and Machine Learning(ML) at the Tiny Edge

Low Latency | **Privacy, IP Protection, Security** | **Offline Mode Operation** | **Bandwidth Constraints** | **Cost Reduction** | **Power Consumption**

- **Key benefits**
  - Ever-increasing demand for integrated solutions
  - Battery powered devices need lower power consumption
  - Small form factor requirements for size constraint devices
  - Data never leaves device: more secure

Sensor



Signal processing (time series low-rate)

- Predictive/Preventative Maintenance
- Bio-signal analysis (healthcare and medical) e.g., pulse detection, EKG
- Cold chain monitoring
- Accelerometer use-cases e.g., fall detection, pedometer, step counting
- Battery monitoring
- Agricultural use-cases e.g., moisture sensing
- Anomaly detection

Audio

Audio pattern matching

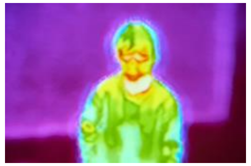
- Security applications e.g., Glass break, scream, shot detection
- Cough detection
- Machine malfunction detection
- Breath monitoring

Voice

Voice commands

- 10 words command set for smart appliance
- Wake-word detection (Always-On voice)
- Smart device voice control
- Voice assistant

Vision



Low-resolution vision

- Wake-up on object detection
- Presence detection
- People counting, people-flow counting
- Movement detection
- Fingerprint

ML Applications at the Tiny Edge with Silicon Labs Series 1, Series 2 Wireless SoCs

Machine Learning Tools for Development in EFR32

SILICON LABS

Software and Tool Support for ML Applications

EDGE IMPULSE

edgeimpulse.com

SensiML™

sensiml.com

Micro.ai

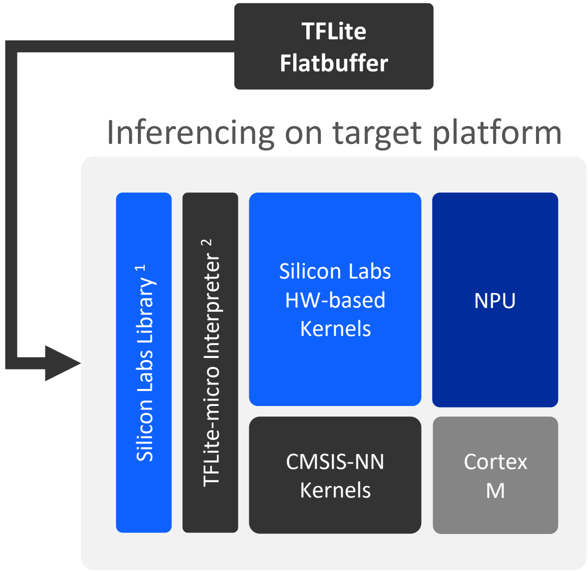
micro.ai

SILICON LABS

Machine Learning Toolkit

[MLTK docs]()

TensorFlow



docs.silabs.com

END-TO-END DEVELOPMENT TOOLS

- 3rd party tools
- End-to-end solutions
- From data gathering to model generation

SILICON LABS MLTK

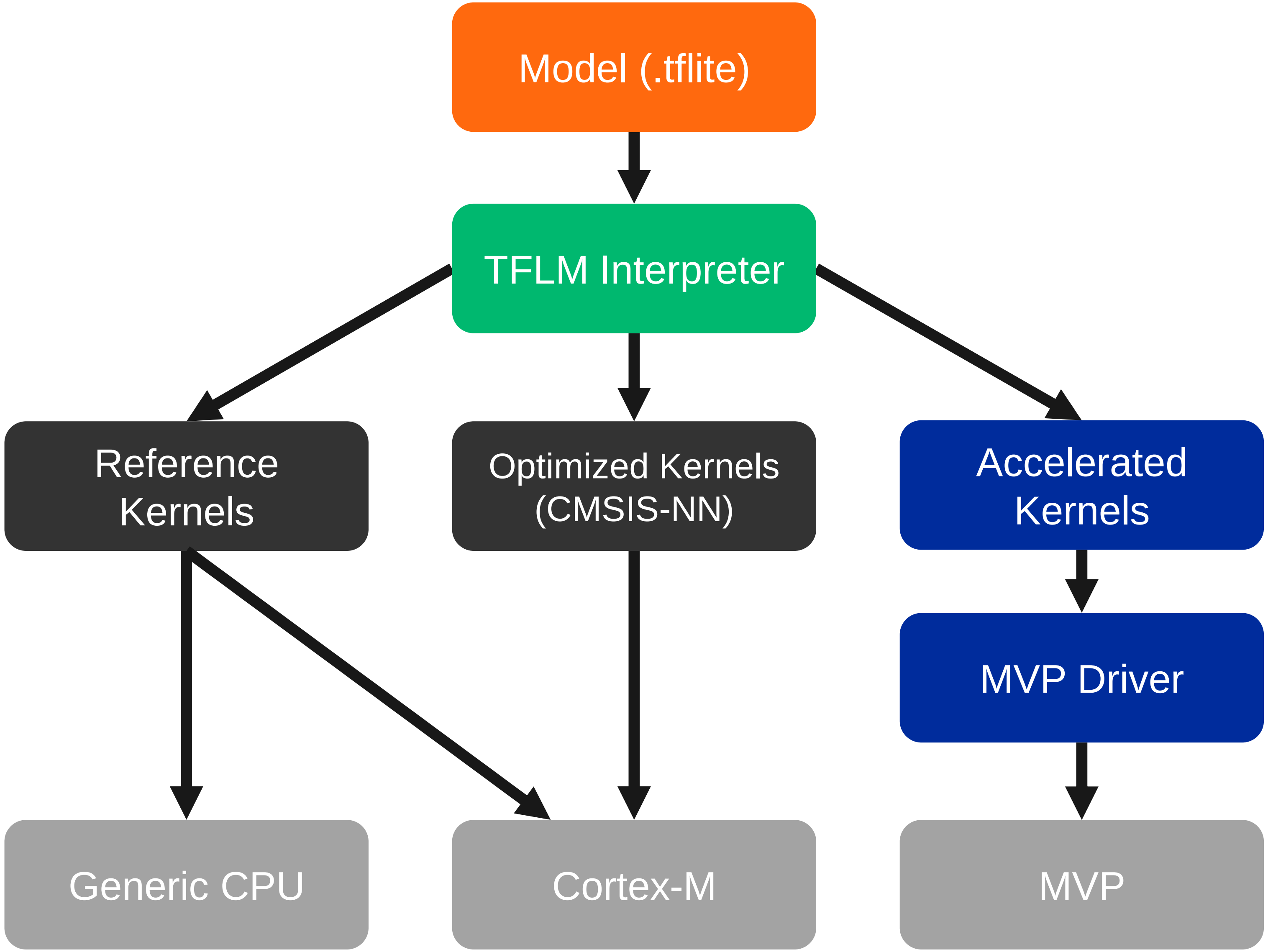
- Python based
- From data processing to model generation
- Tools for model performance evaluation
- Preprocessing tools equivalent to those in GSDK

TENSORFLOW LITE MICRO

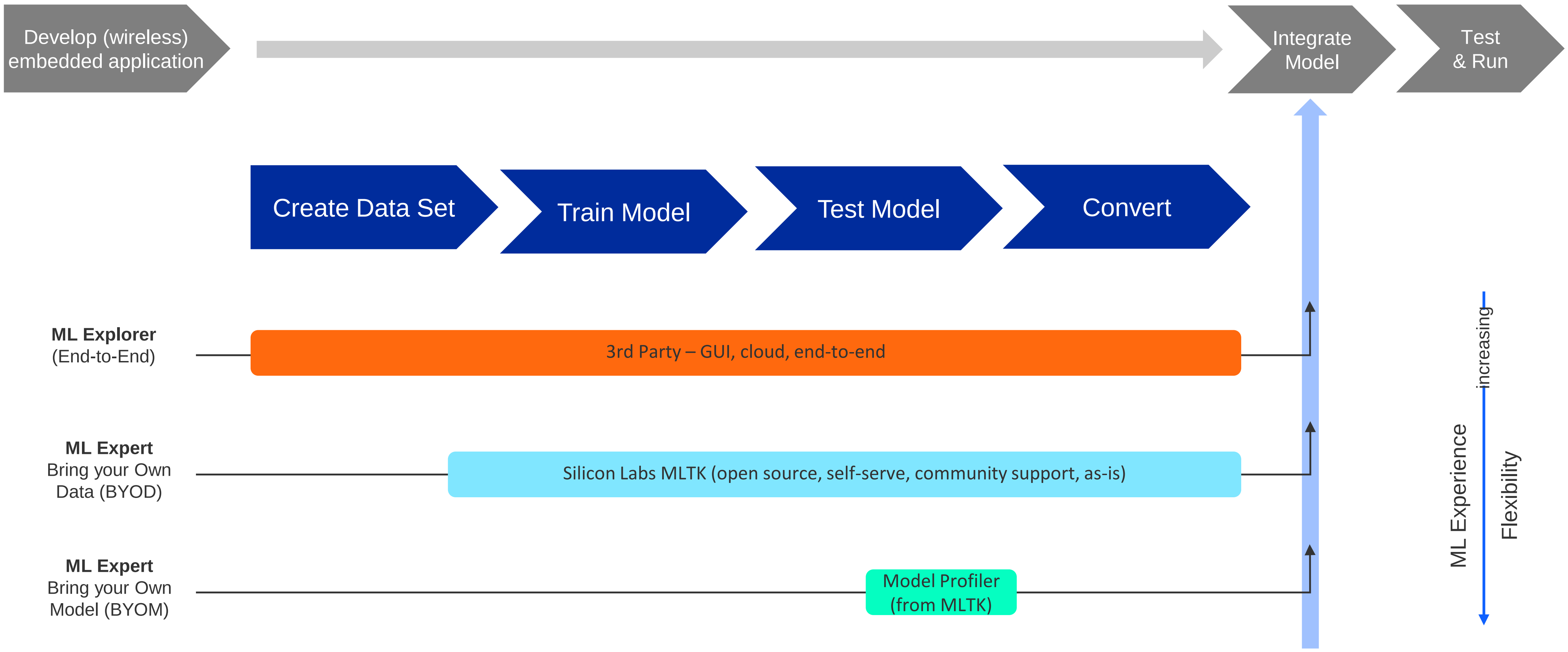
- GSDK integration of TensorFlow lite for microcontrollers
- Reference, optimized and acceleration kernels
- Model should be created offline

TensorFlow Lite Micro in the Gecko SDK

- **C-array generated from model flatbuffer (.tflite)**
- **Interpreted at run-time**
- **Kernels execute each layer of the model graph**
- **3 layers of kernels**
  - Reference kernels
    - CPU agnostic code
    - Slow
  - Optimized kernels
    - Tailored for CPU
    - E.g., using ARM SIMD instructions
    - Faster
  - Accelerated kernels
    - E.g., running on the xG24 Matrix Vector Processor
    - Fastest (2-3 times faster than optimized kernels)



Embedded Development with Machine Learning (Supervised)

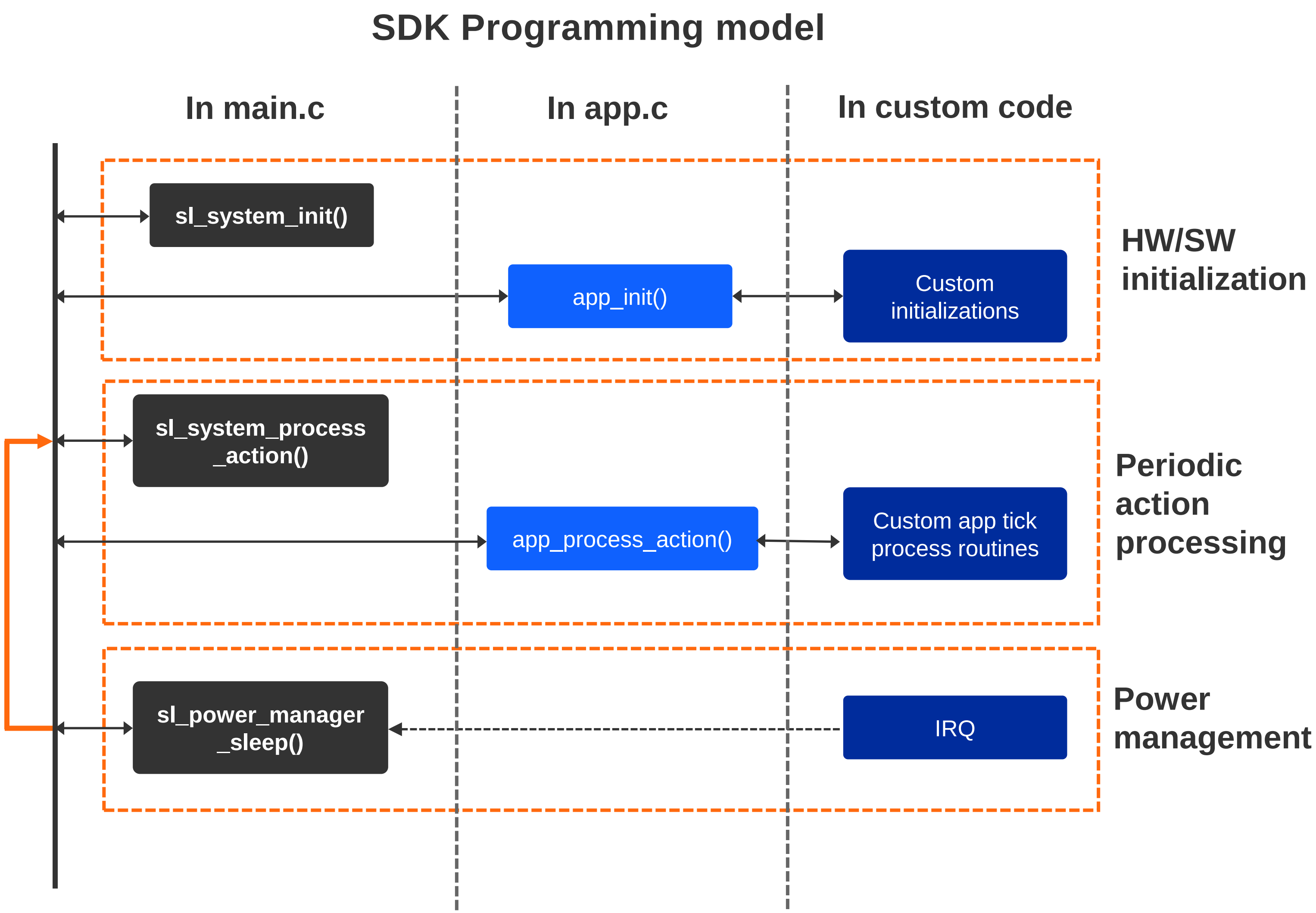


There are 3 workflow options depending on level of Machine Learning experience, and implementation flexibility desired

Simplicity Studio v5

Programming Model and Software Components

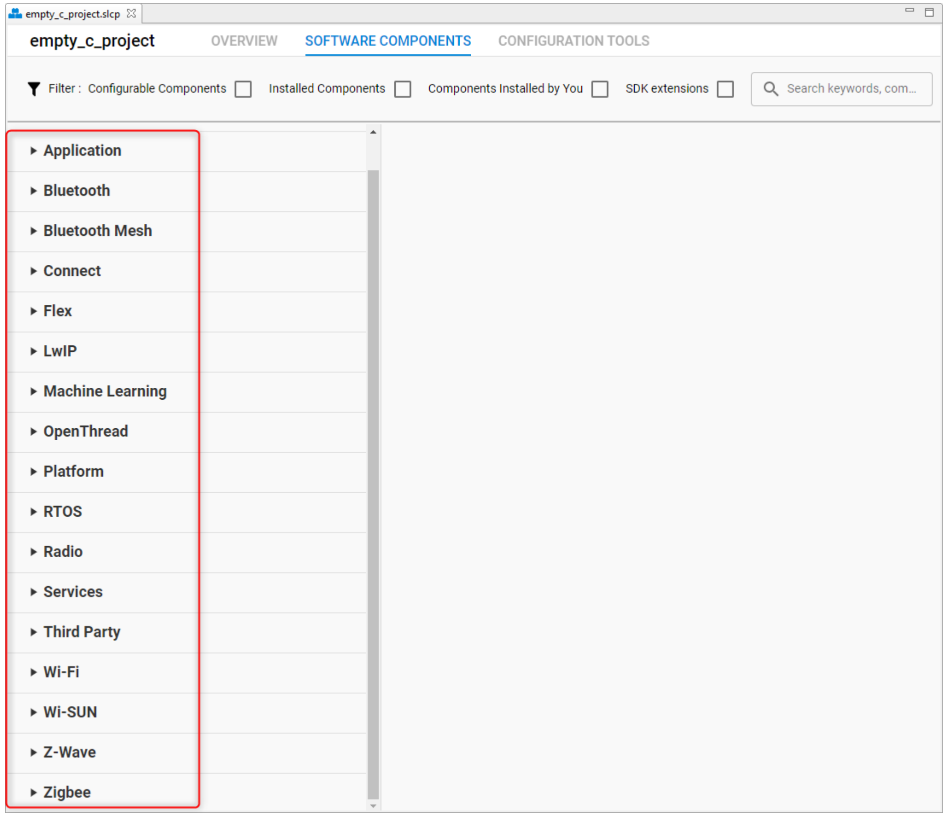
SSv5 Architecture Concepts and Programming Model



- **Gecko Platform**
  - Software catalogue
  - Presented as software components
- **Project Configurator**
  - Manage software components and access to other GUI tools
    - Component Editor
    - Pin Tool
    - GATT Configurator
    - Radio Configurator

SSv5 Software Components

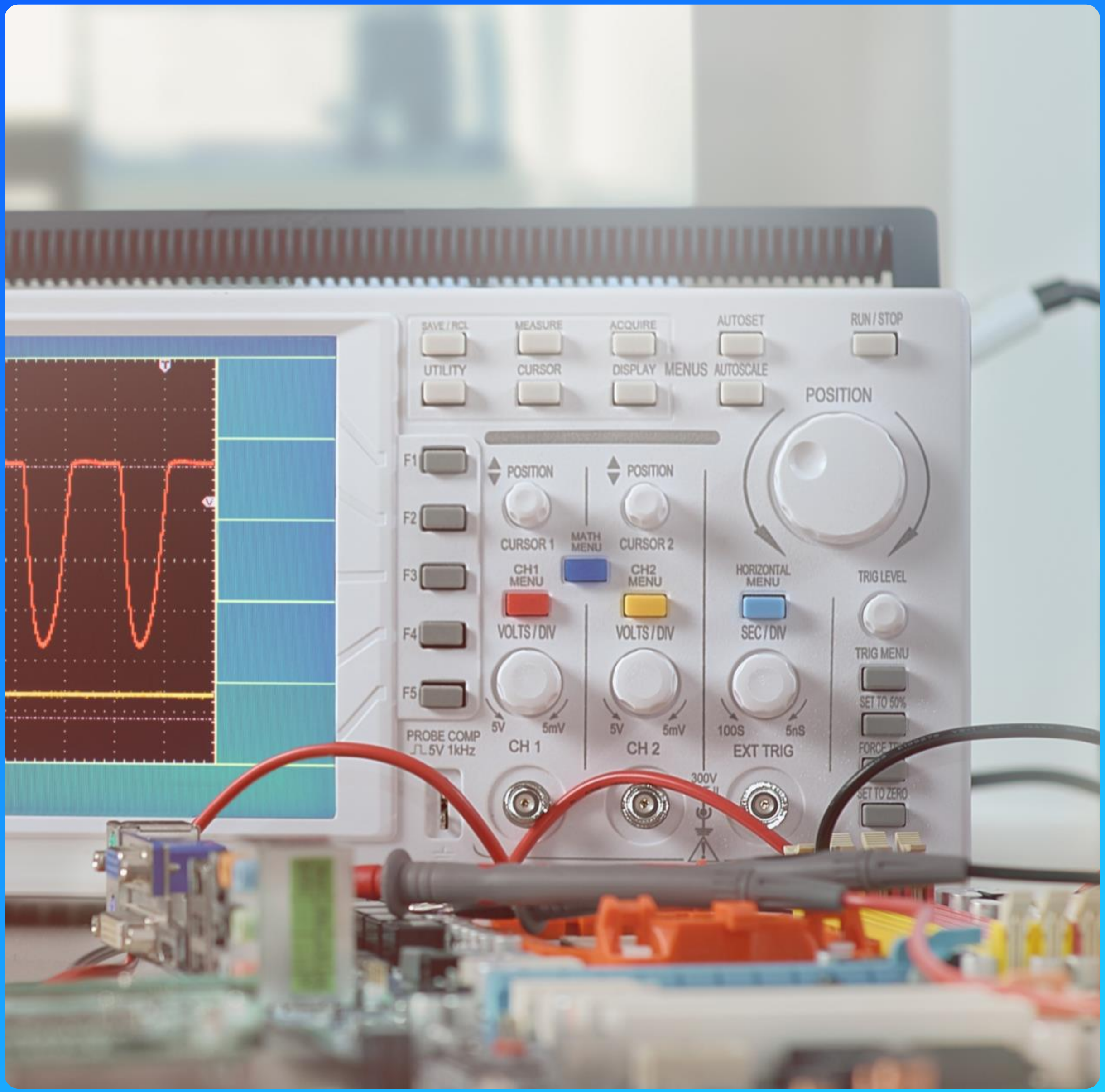
- **Software Components**
  - Added through the Project configurator
  - Install, uninstall, configure and instantiate
  - Automatically add source, header and autogen code



- **External Device Drivers (multiple)**
  - APIs for Silabs and 3rd party devices
  - Development kit sensors
- **Board Control**
  - Enable/disable development kit features
- **Sleep Timer**
  - Software timers based on low-frequency hardware timers
- **Power Manager**
  - Mange the system energy modes

- **Detailed information**
  - docs.silabs.com – Gecko Platform

Lab session


AUTOSET
RUN / STOP
MEASURE
ACQUIRE
UTILITY
CURSOR
DISPLAY
MENUS
AUTOSCALE
POSITION
F1
F2
F3
F4
F5
POSITION
POSITION
CURSOR 1
MATH MENU
CURSOR 2
CH1 MENU
CH2 MENU
HORIZONTAL MENU
TRIG LEVEL
VOLTS / DIV
VOLTS / DIV
SEC / DIV
TRIG MENU
SET TO 50%
5V
5mV
5V
5mV
100S
5nS
PROBE COMP
5V 1kHz
CH 1
CH 2
EXT TRIG
SET TO ZERO

Lab Introduction

- **Objective:** Create a bare-metal application for sensor data collection using the xG24 dev kit, Simplicity Studio v5 and the Gecko Platform.

- **Requirements**
  - Hardware:
    - EFR32xG24 Dev Kit xG24-DK2601B (BRD2601B)
    - Micro-USB to USB Type-A cable
    - **(Optional)** WSTK BRD4001A or WSTK PRO BRD4002A
  - Software
    - Simplicity Studio v5
    - Gecko SDK Suite 4.0.2 or later
    - Accompanying lab source code

Lab Sections

STAGE 1

Creating and configuring a bare-metal project

STAGE 2

Integrating the custom application code

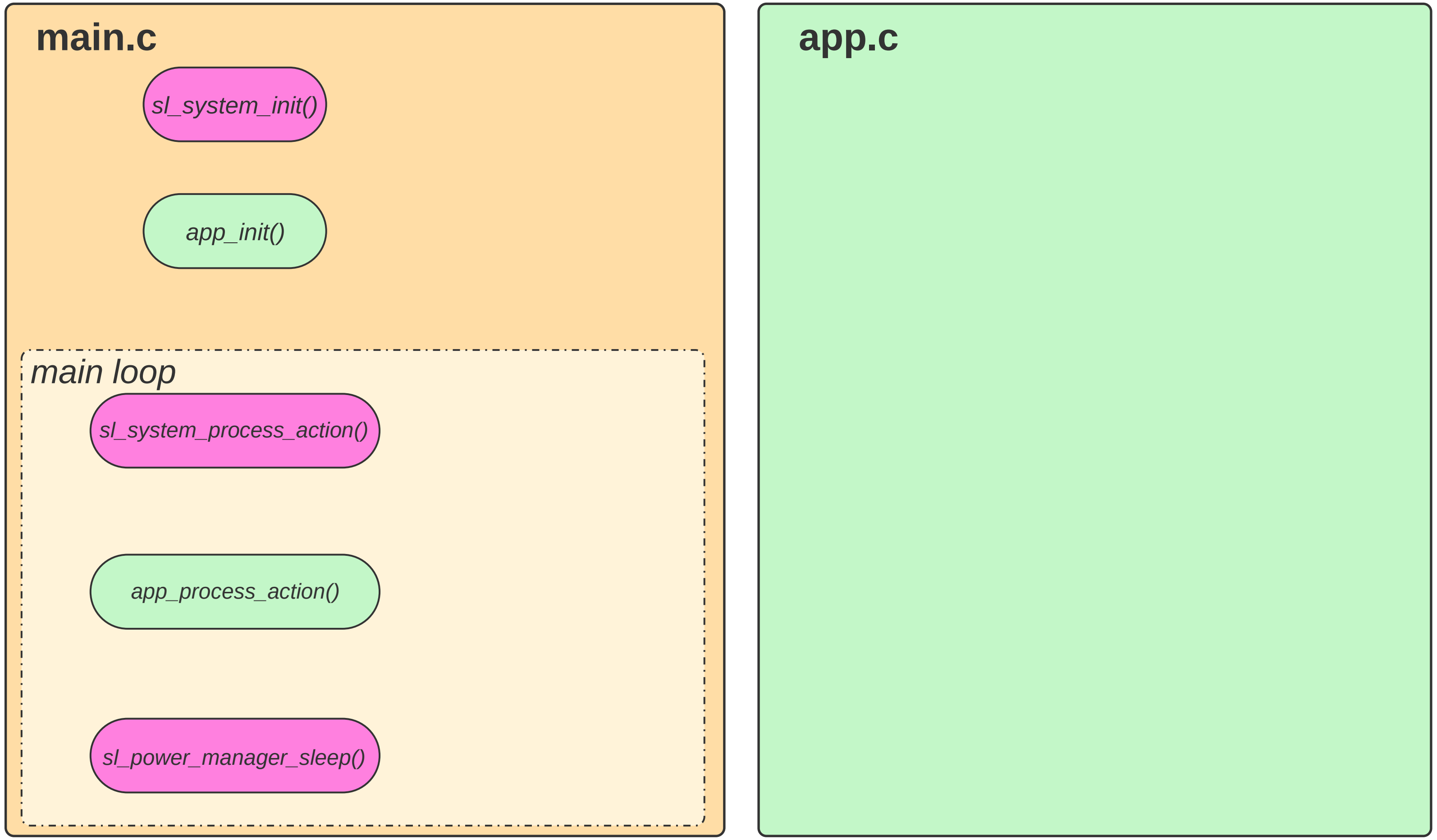
STAGE 3

Interpreting the application code

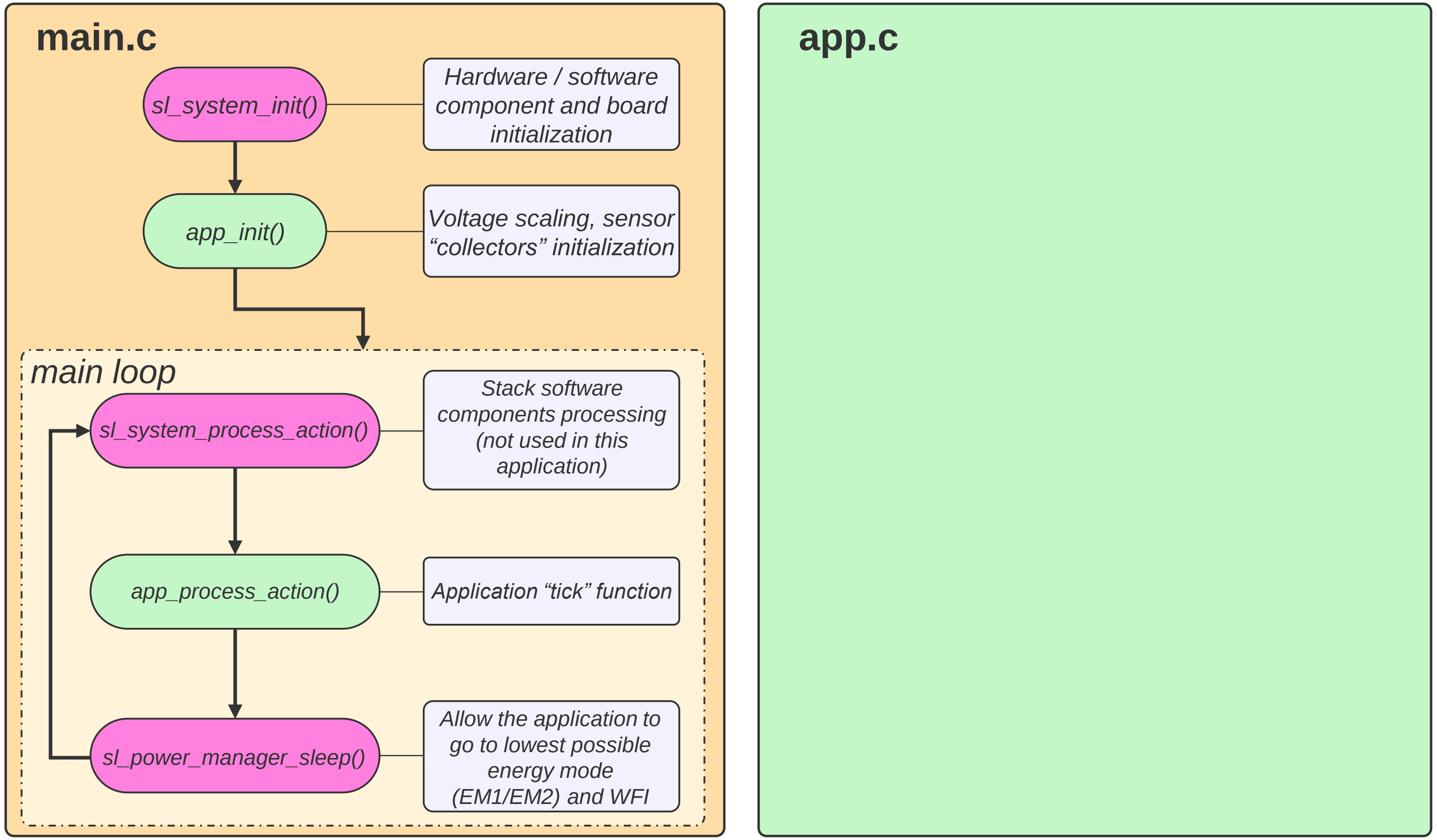
STAGE 4

Testing the application and energy profiler measurements

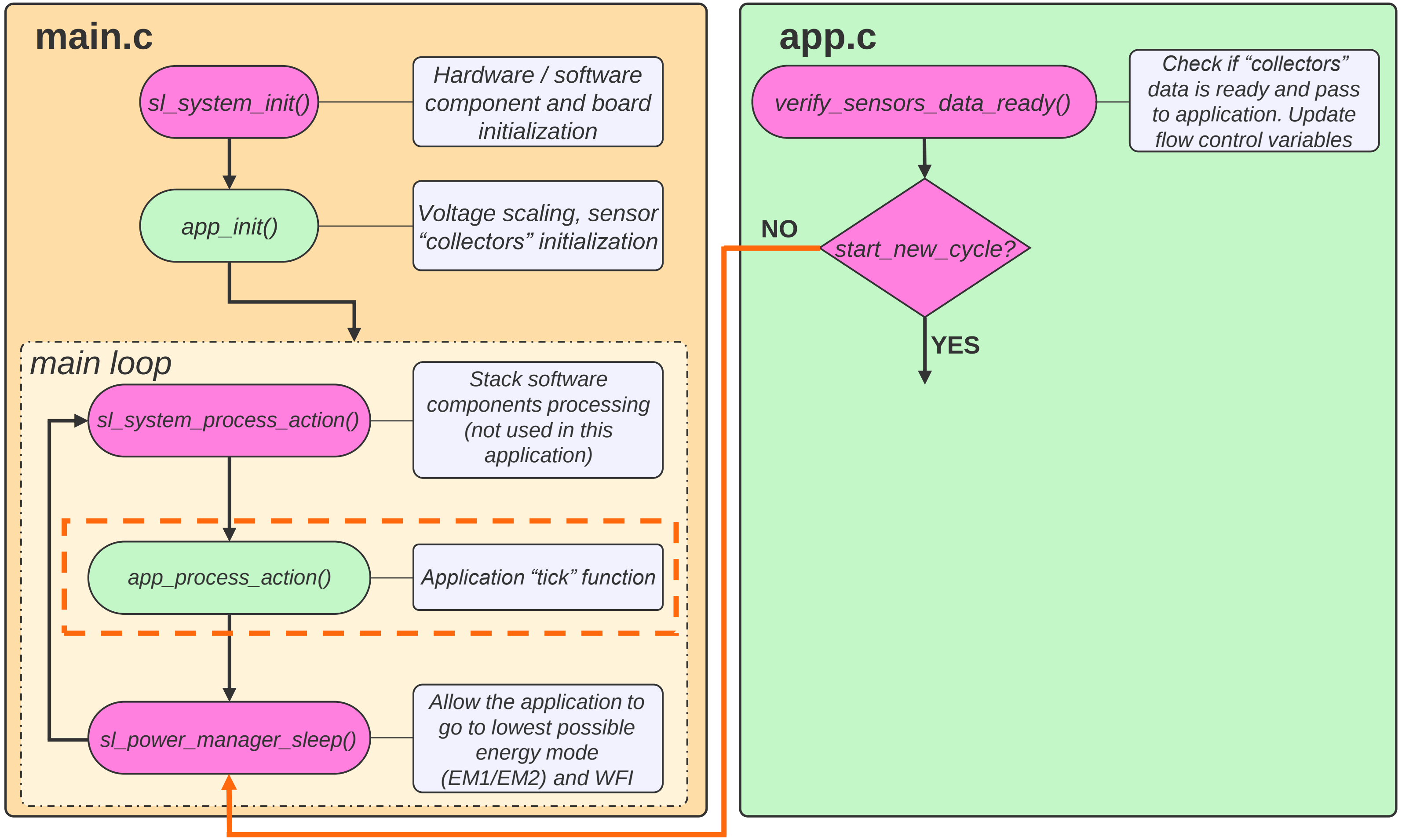
Interpreting the Application



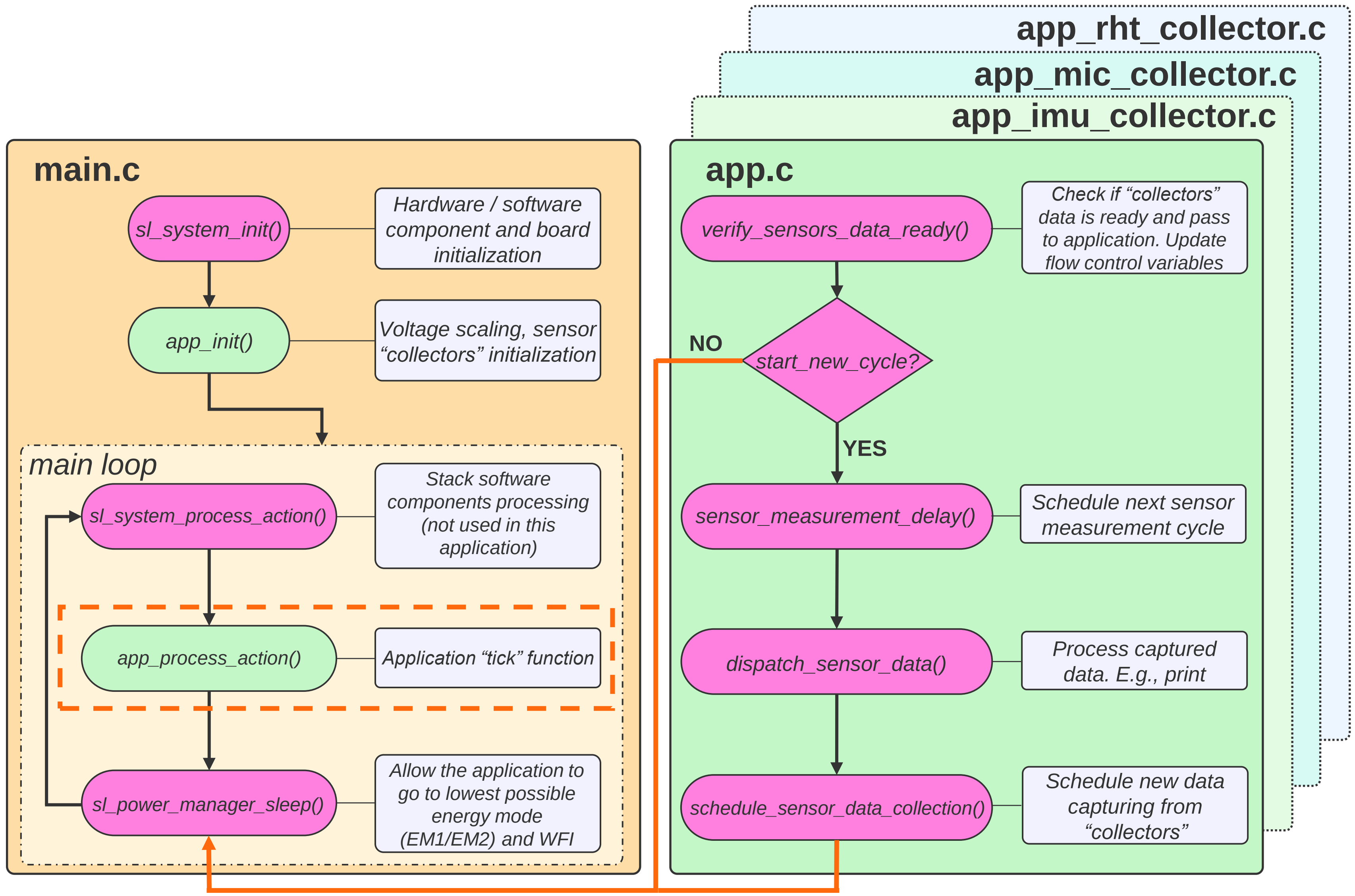
Interpreting the Application



Interpreting the Application



Interpreting the Application



Closing thoughts

- **Energy profiler**
  - In data gathering applications consider the following 5 factors:
    - Latency
    - Number of samples collected
    - Sampling frequency
    - Number of sensor sources
    - Data processing
  - Finish early, sleep early

- **Application limitations:**
  - No check to verify if the data was fully collected
  - No timed method to acquire RHT data
    - Could be implemented with sleep timer
  - Software components are not 100% efficient, e.g:
    - I2S microphone: Captures and discards 4096 samples BEFORE getting the requested samples
    - IO STREAM: Interrupt-based instead of LDMA based
    - I2CSPM: Polled vs interrupt or LDMA
  - Powering on-board sensors requires GPIO
    - Slight consumption increase in EM2
  - I2C sensors are powered through the same enable signal