tinyML® EMEA

Enabling Ultra-low Power Machine Learning at the Edge

June 26 - 28, 2023

TINY ML

www.tinyML.org

28 June 2023

Data Pre-processing on Sensor Nodes for Predictive Maintenance

POLYN TECHNOLOGY
NEUROMORPHIC ANALOG SIGNAL PROCESSING

Aleksandr Timofeev, CEO

PROBLEM: EXPONENTIAL GROWTH OF SENSOR DATA THREATENS TO OVERWHELM RESOURCES OF HUMANITY

Electronic Sensors

Use of deterministic operations
is a barrier for energy efficiency

- **Sensor technologies** are experiencing exponential growth with forecasts of **~45 trillion sensors in 2032** that will generate >1 million zettabytes (10^{27} bytes) of data per year[1]
- **Sensor raw data** processing is a **bottleneck**

Bio Senses

True parallel data processing is the key
for unprecedented energy efficiency

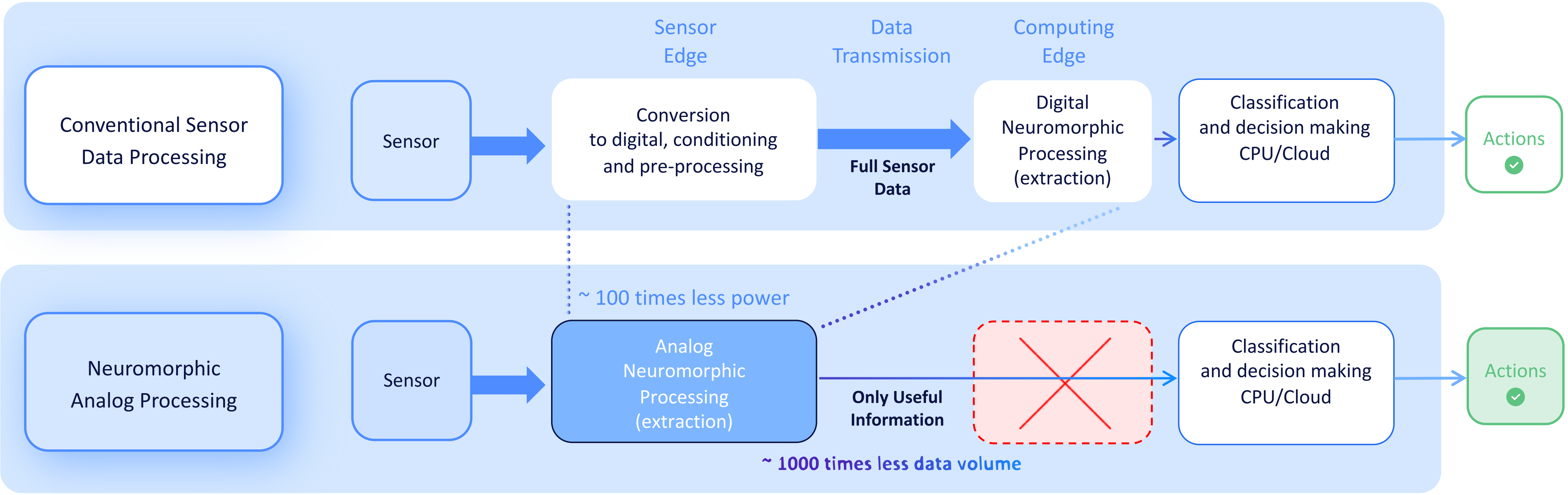
Biological systems are **100,000 times more energy efficient**[2]
than digital computer systems

(1) https://www.frontiersin.org/research-topics/31093/nanoelectronic-devices-for-analog-hardware

(2) Landauer's principle defines the theoretical limit of energy consumption of computation

SOLUTION: NEUROMORPHIC ANALOG SIGNAL PROCESSING OF SENSOR RAW DATA

- Analog neuromorphic cores based on true parallel process extract useful information from sensor raw data more efficiently than traditional digital methods based on deterministic operations
- Extracting useful information from sensor raw data is based on mimicking biological sensing processes of nature



- An analog neuromorphic front-end with a neural network implemented in analog on the chip is several orders of magnitude more efficient than a digital system deterministic calculations of a math model

POLYN TECHNOLOGY

TECHNOLOGY PLATFORM TO MAKE NFEs

- Neuromorphic Analog Signal Processing technology platform is the key to developing and manufacturing the neuromorphic front end
- NASP technology allows conversion of trained neural networks into an analog neuromorphic core and its manufacturing

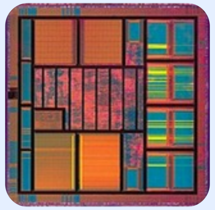
Select Trained NN

Trained Neural Net math model in any standard format: Keras, TensorFlow or others

Assisted and supported by **POLYN**

Convert Math Model Into Chip Layout

IP



Math model generated with NASP Compiler

D-MVP – Product digital simulation for validation

Chip model Netlist ready

Fully automated process by **POLYN**

NASP GDSII files ready for any foundry: TSMC, Global Foundries, Samsung, etc.

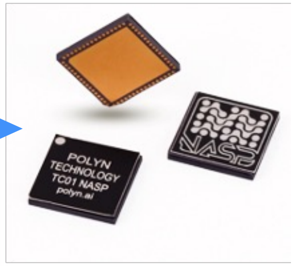
Fully automated process by **POLYN**

Chip Production at a Fab

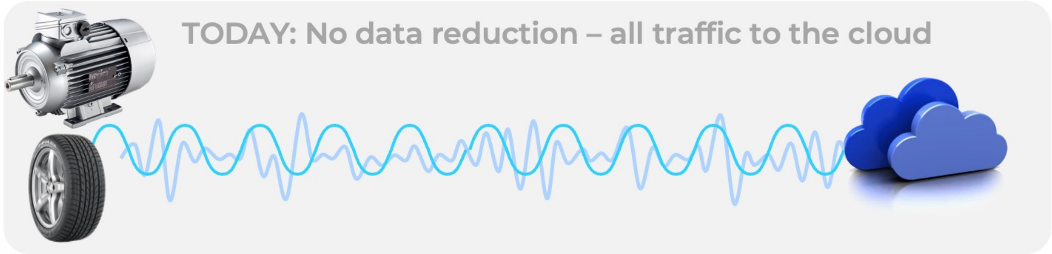
NASP chips manufactured from standard analog components according to NASP neuron design

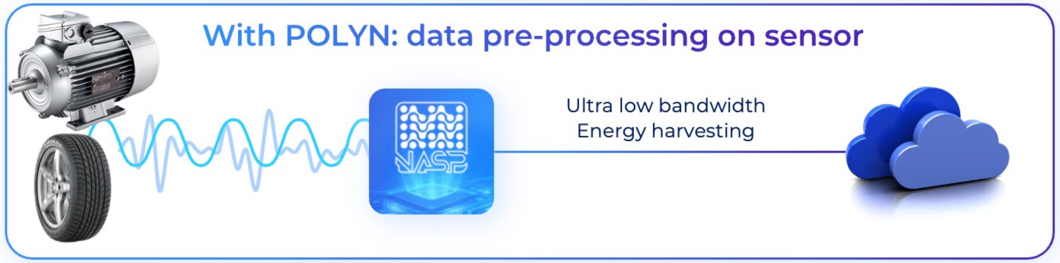
Any Fab with 40-90 nm process

Chips for the Sensor Node Product



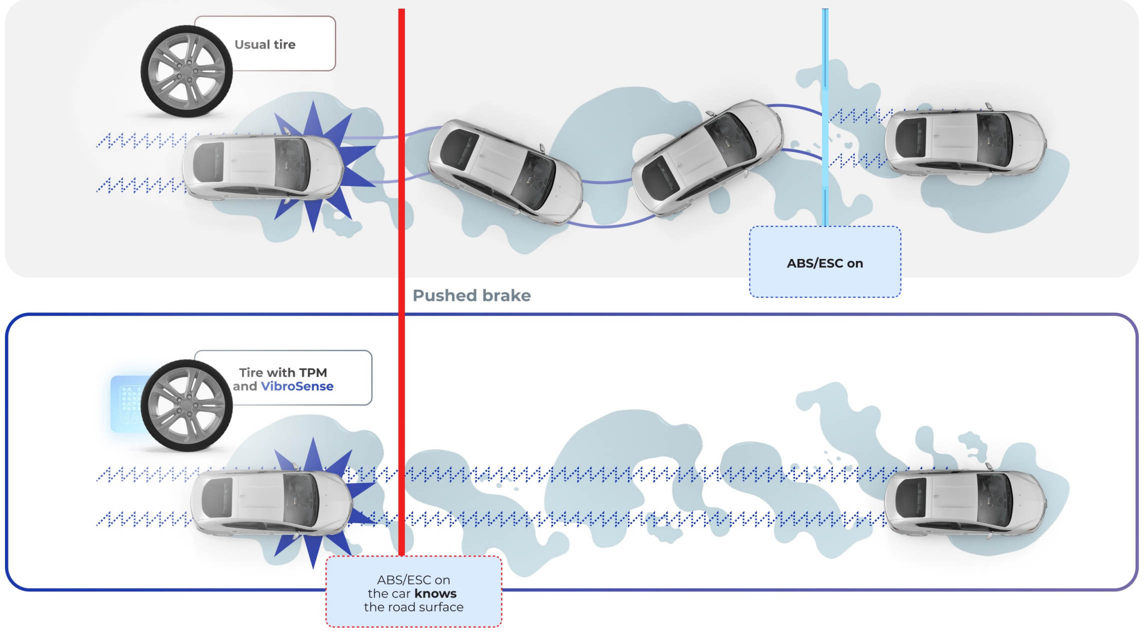
APPLICATION: VIBRATION MONITORING



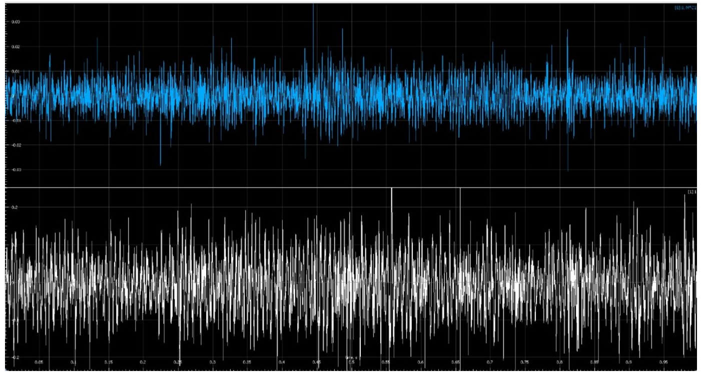


NFE enables data volume reduction by more than 1000 times, and power consumption reduction by 100 times to support ultra-narrow band communication and very long battery life or energy harvesting based designs

Example: embedded Temperature Pressure Monitoring (TPM) Sensor with integrated 3-axis accelerometer and VibroSense chip.
Proof-of-Concept (PoC)
with a **tire manufacturer**



CHALLENGES OF VIBRATION (ACCELEROMETR) SIGNAL PROCESSING



Input signal from 1 axis accelerometer >0,6 Mbps for 10 kHz bandwidth

- Data compression by a neural network of efficient size
- Avoid lossy compression
- Necessary to find an encoder function to minimize the information loss or maximize mutual information:

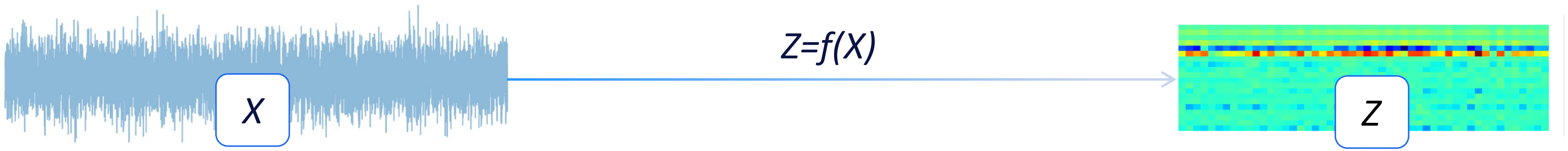
$$\hat{f} = \underset{f \in F}{\operatorname{argmax}}\, I(X; f(X))$$

where ***X*** is an input signal,
f is an encoder,
I is mutual information.

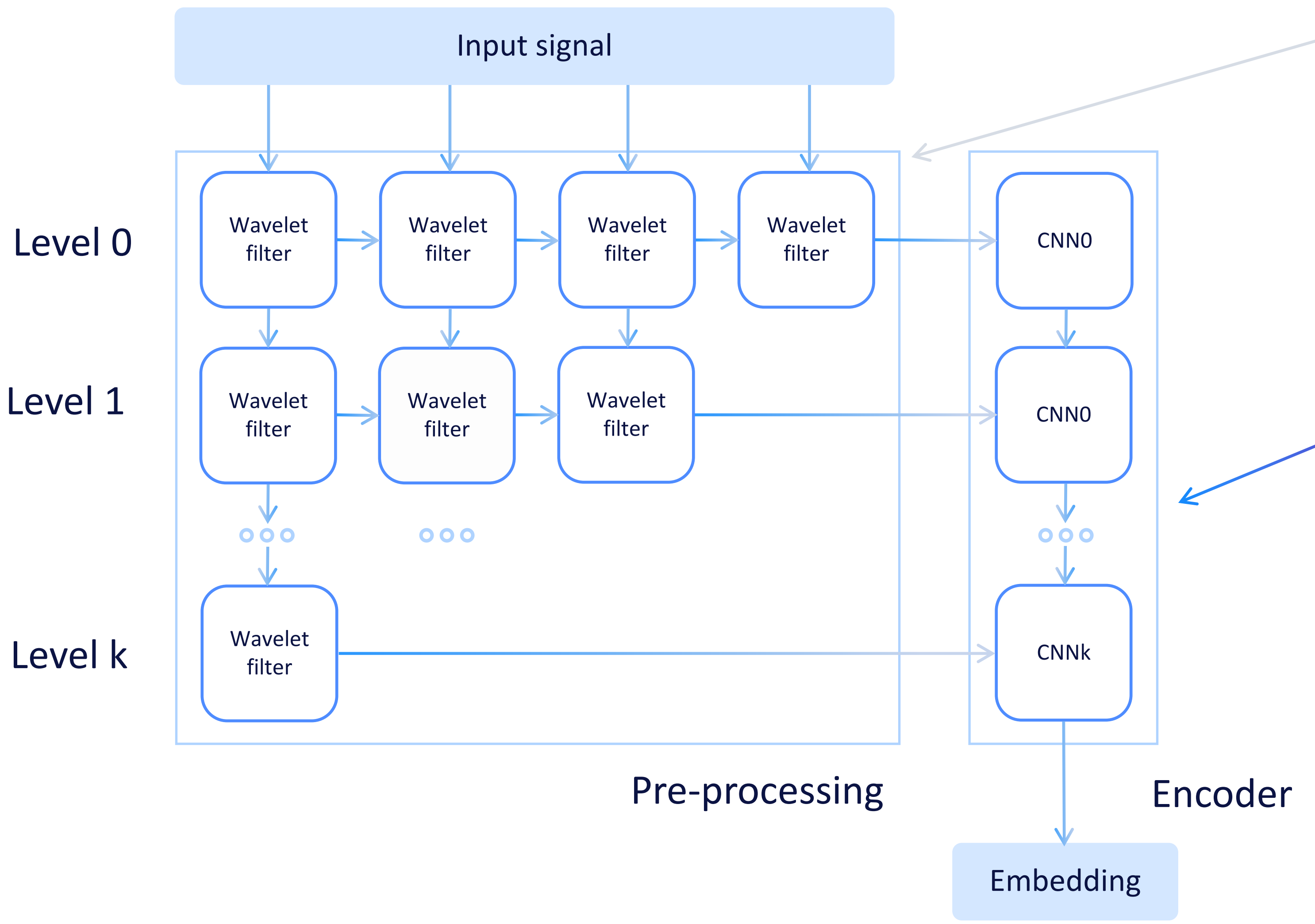
The above is hard to optimize directly. We can minimize the autoencoder loss by searching for a decoder function ***g*** as well:

$$\hat{f}, \hat{g} = \underset{f \in F, g \in G}{\operatorname{argmin}}\, L(X, g(f(X)))$$

Visualization of the encoding process: transforming input signal into latent representation



OVERVIEW OF THE NEURAL NETWORK MODEL ARHITECTURE



WHY WAVELETS:

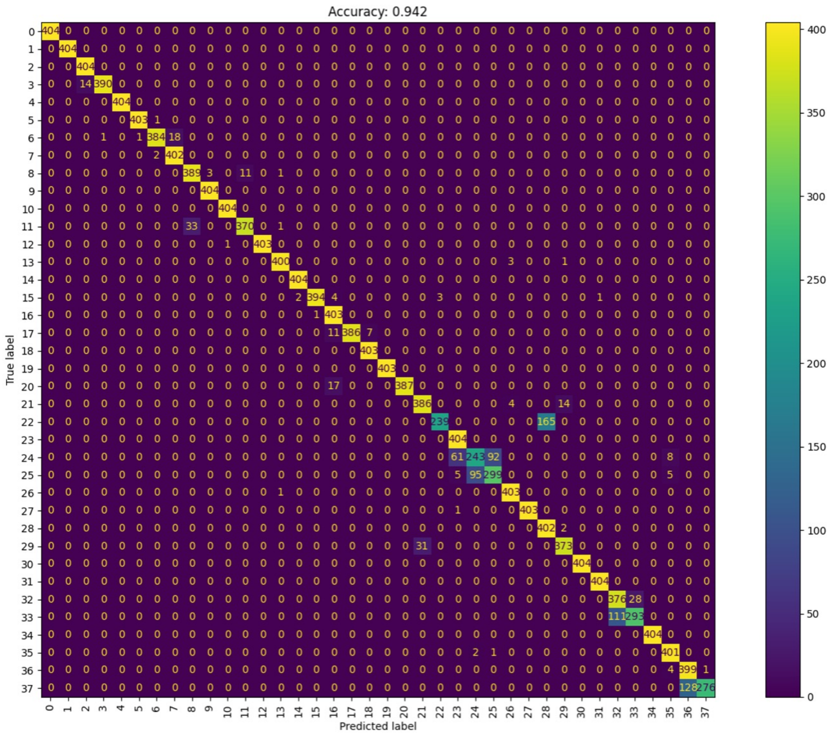
- better time-frequence resolution balance than FFTs
- good fit to the chip: short shared convolutions

AUTOENCODER VARIANTS:

- **Unrestricted:** highest compression rate but more entangled latent space
- **Latent space regularization (VAE):** lower compression rate but 'nicer' latent space

METRICS AND RESULTS OF TEST DATA SET

- 38 classes (balanced)
- 11 hours
- 921 millions of data points

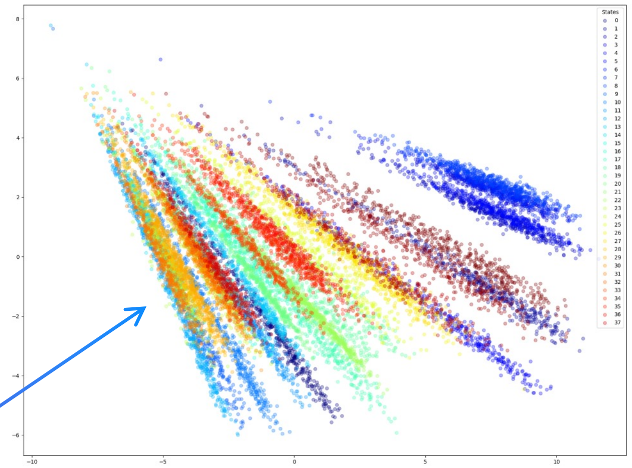


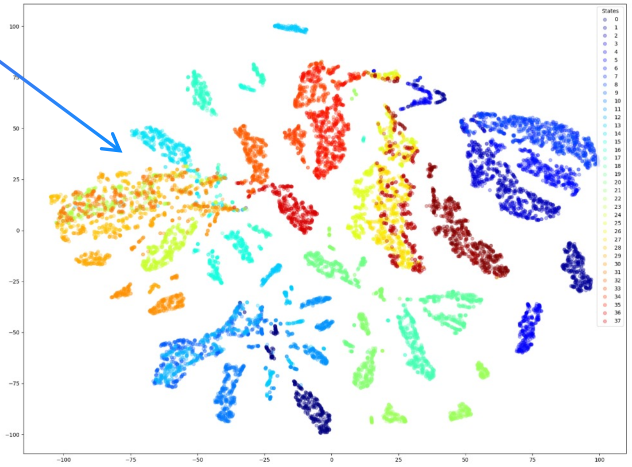
Confusion Matrix. Accuracy: 0.942

PCA projection

Well-structured low-dimensional space

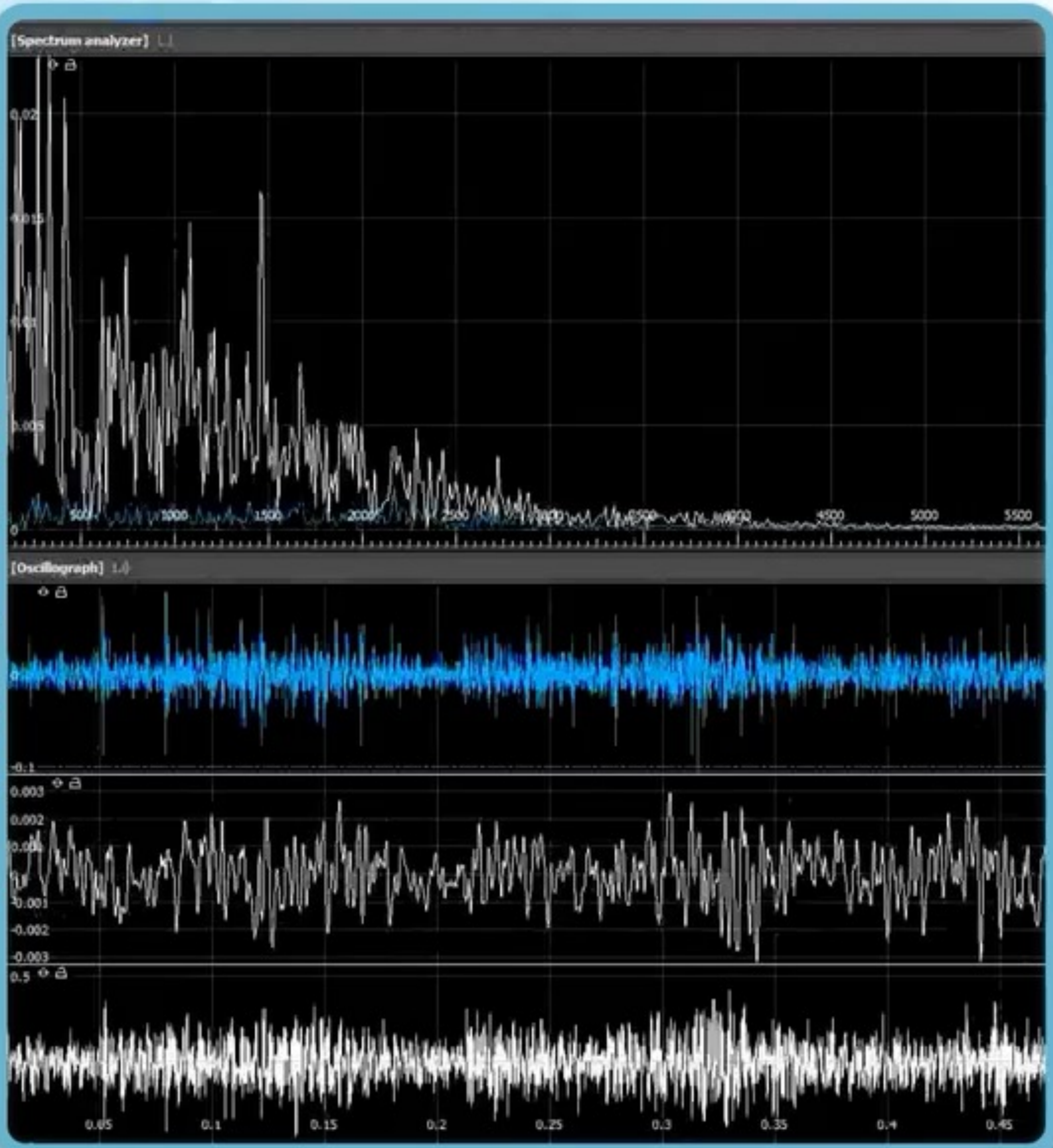
T-SNE projection



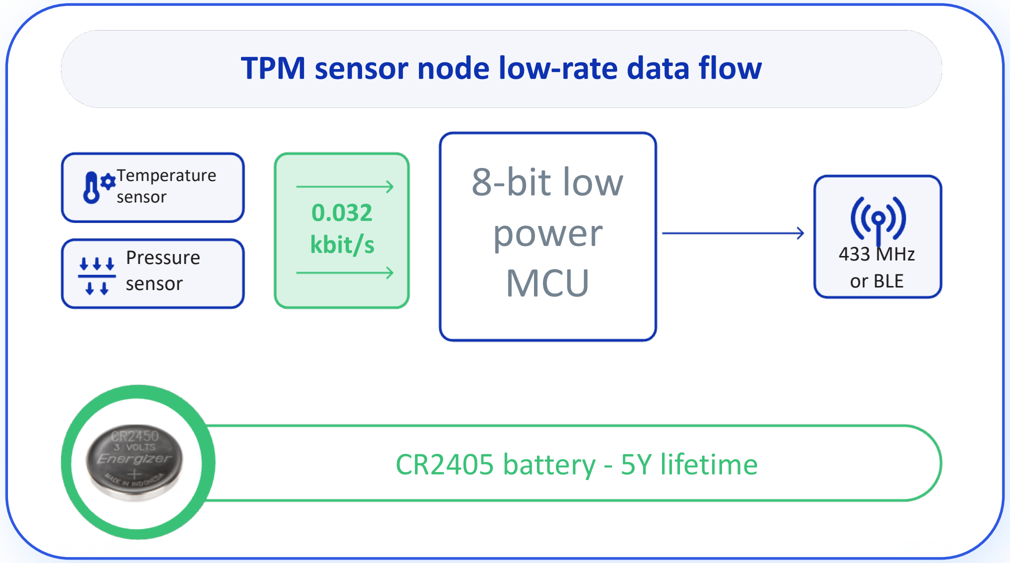


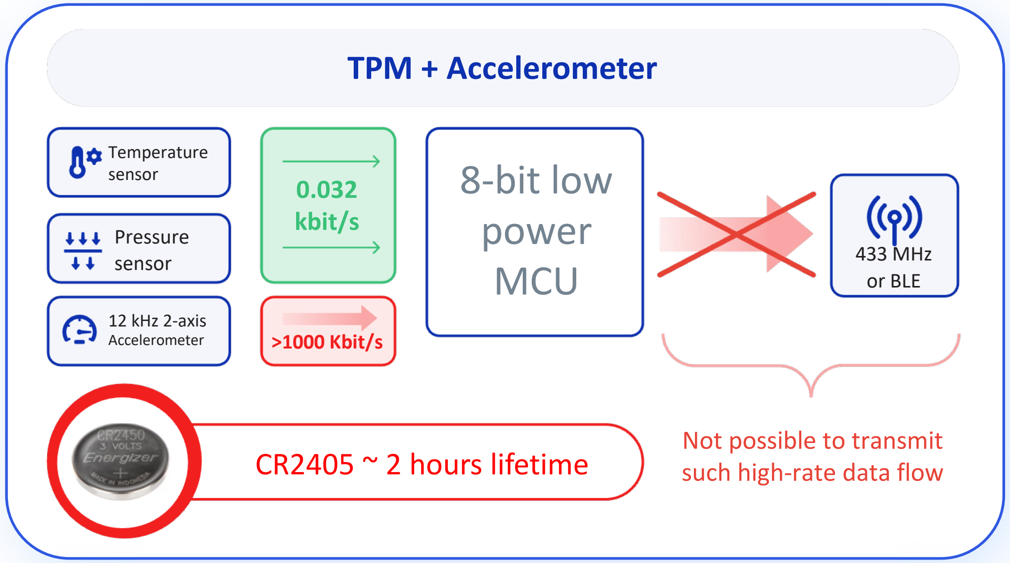
1 axis, 24 kHz sampling rate, 24 bit ADC

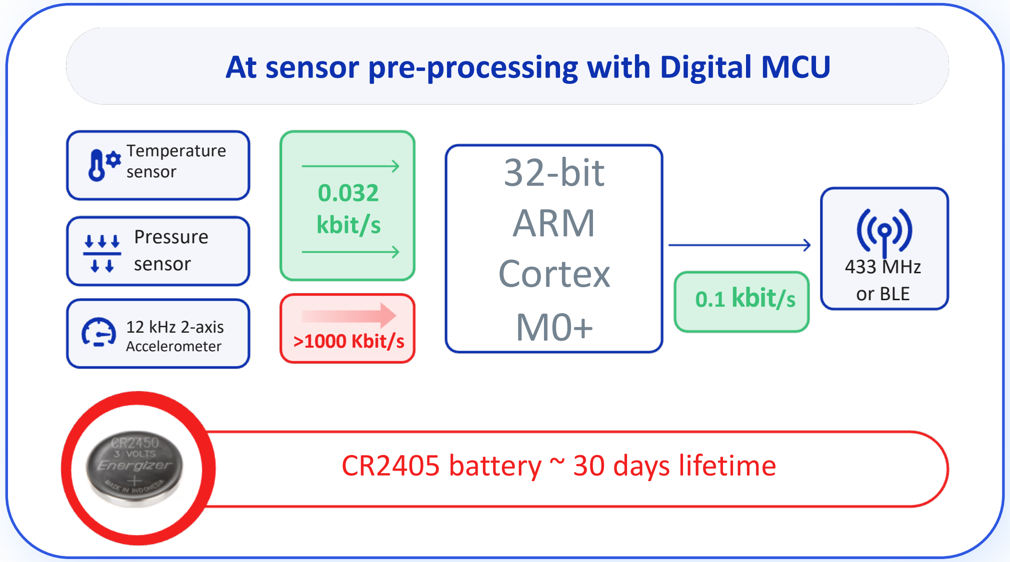
Min
562,5 Kbit/s

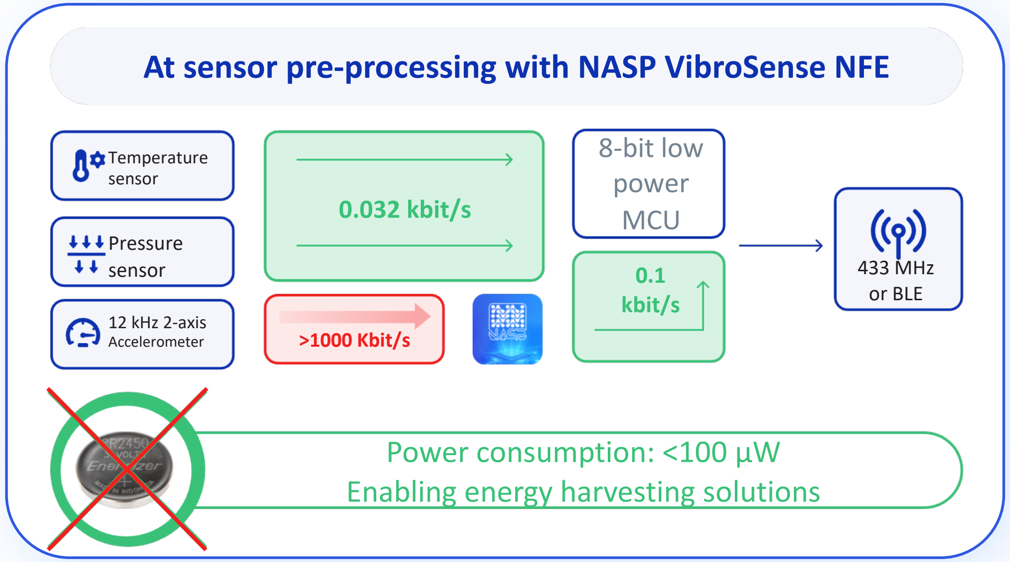


NFE – ENABLING SOLUTION



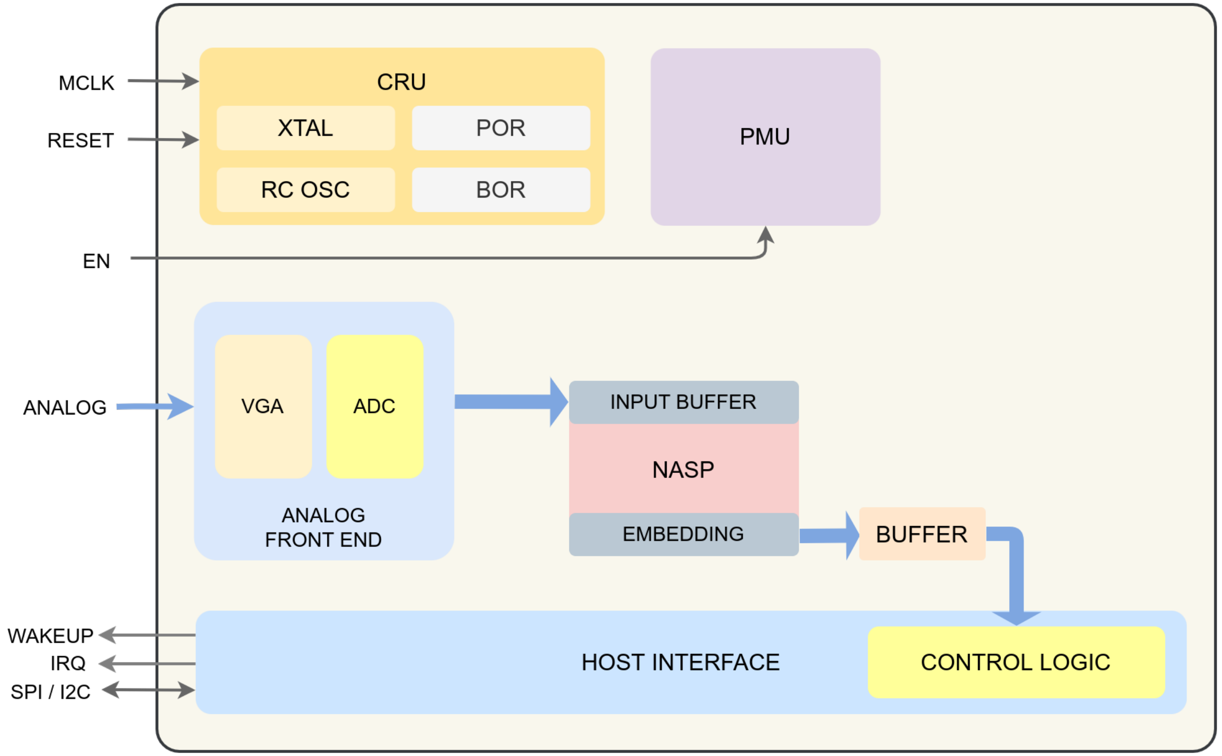






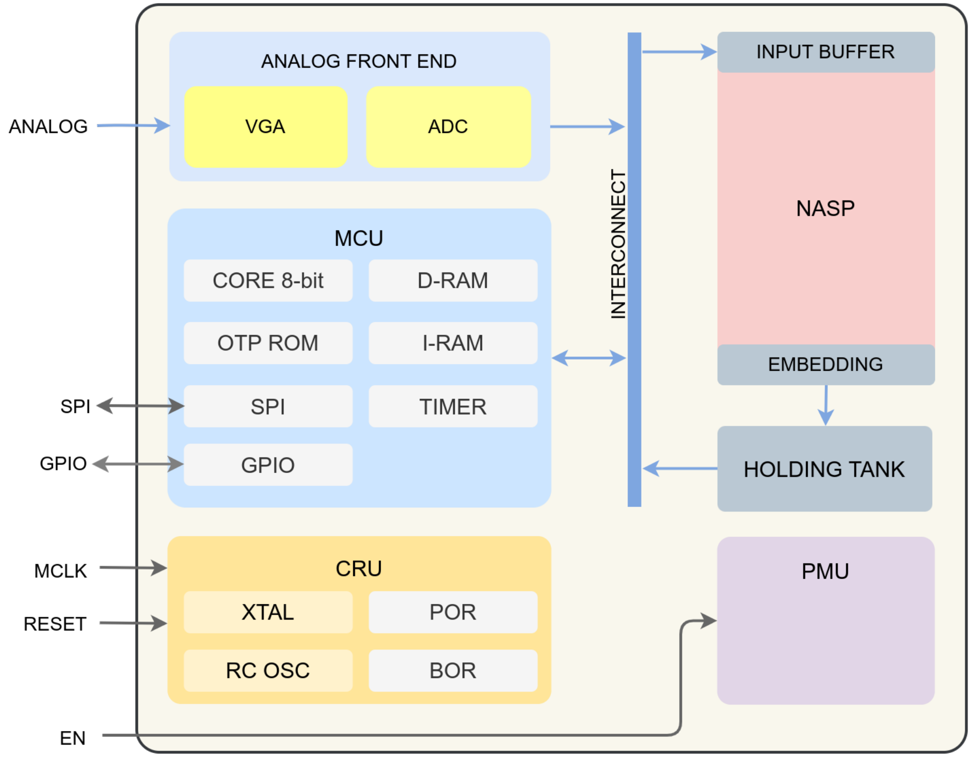
VibroSense^{tm} CHIP ARCHITECTURE

Light version – Host control



Disabled mode: Less than **10 nW**
Enabling mode: Less than **100 µW**
Configuration and control - external Host MCU
Seamless integration with existing sensor node designs, such as standard TPMU sensors with integrated 8-bit MCU

Full version – integrated 8-bit ultra low power MCU



Disabled mode with RTC: Less than **500 nW**
Enabling mode: Less than **1 mW**
Capability to function as a host for LoRa or other RF modules
Ideally suited for industrial IoT modules

NFE FOR A VARIETY OF APPLICATIONS

NFE is a new technology based on Neuromorphic Analog computation on-chip that can be applied to various tasks

IIoT sensors - long battery life

INDUSTRY 4.0

Voice Processing

VOICE PROCESSING

Always-ON health monitoring

WEARABLES

Eye Tracking for AR/VR

XR and LOW-RESOLUTION IMAGES

PMSM - Motor control

ELECTRICAL MOTORS

Smart Pallets
Smart Tires

SMART IOT

We invite anyone with a model for a sensor:
NASP will convert it into a chip, efficient, sustainable for next generation of IoT devices

Thank You!

info@polyn.ai

POLYN TECHNOLOGY
NEUROMORPHIC ANALOG SIGNAL PROCESSING

Copyright Notice