Akraino Use Case: Robotics

LF Edge at OCP Summit 2023
Akraino Summit 2023

Contents

- **Organizations Participating**
- **LF Edge and Akraino**
- **Challenges & Motivation**
- **Use Case Description**
- **Achievements & Results**
- **About Akraino**

Organizations Participating

- **Ritsumeikan Univ**
  - comprehensive university with 18 faculties and 20 graduate schools
  - students study social sciences, humanities and natural sciences
  - 3rd largest number of students in Japan
- **Fujitsu**
  - information & communications technology equipment and services firm
  - providing IT & IT infrastructure and other services
- **Signalogic**
  - solutions for high capacity, high reliability media
  - deployed in HPC, telecom, lawful intercept, deep learning, and applications

LF Edge & Akraino

- **Akraino**
  - LF Edge open source community creating application blueprints for Edge Computing
  - spanning edge computing use cases in 5G, AI, Edge IaaS/PaaS, IoT, and robotics
  - for both provider and enterprise edge domains
- **Cyber-Physical Systems (CPS) Akraino Blueprints**
  - Robot Basic Architecture
    - based on Ritsumeikan and Fujitsu SSES (Sensor-rich Soft End-effector System)
  - Smart Data Transaction
    - for cloud management and orchestration of robotics systems
- **Synergies with LF Edge communities**
  - EdgeX Foundry
    - enables nearby IoT gateways to allow robot communication via WiFi or Bluetooth
  - EVE
    - collect and transmit 9-axis sensor data to cloud algorithms for failure prediction
    - robots require continuous monitoring and maintenance

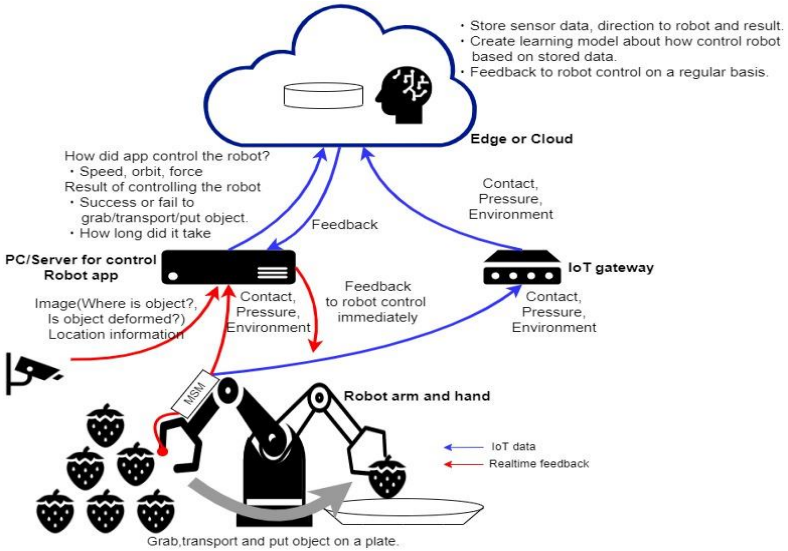
Challenges & Motivation

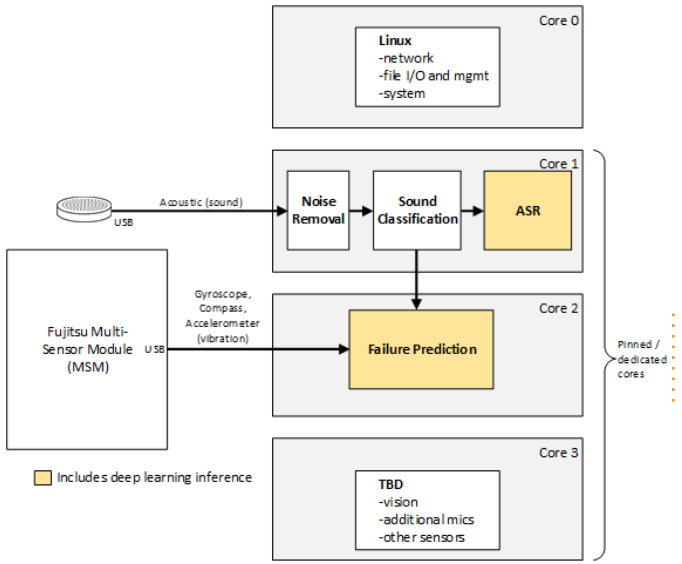
- **Industry 5.0**
  - massive, automated customization
  - human + robot “co-working”, expanding from conventional assembly line to agriculture, food preparation, and other mfg with high degree of human interaction
  - achieving Industry 5.0 objectives faces key challenges …
- **Non-uniform objects**
  - robots must handle objects with diverse shapes, frictional properties, and range of weight and thickness
  - widely variable environment due to wetness, clutter, customers, etc.
- **Human interaction**
  - robots must communicate quickly and reliably with people
  - human safety is paramount, robots must handle emergency voice commands
    - large vocabulary is required – commands are unpredictable in wording, sequence, and level of stress (emotion)
  - cannot be cloud dependent – no dependencies on WiFi or Internet connectivity

Use Case Description

- **Sensor + robot fusion**
  - touch and tactile data are collected using multiple sensors
  - analyzed by cloud AI algorithms
  - fed back to the robot to achieve delicate handling according to each object

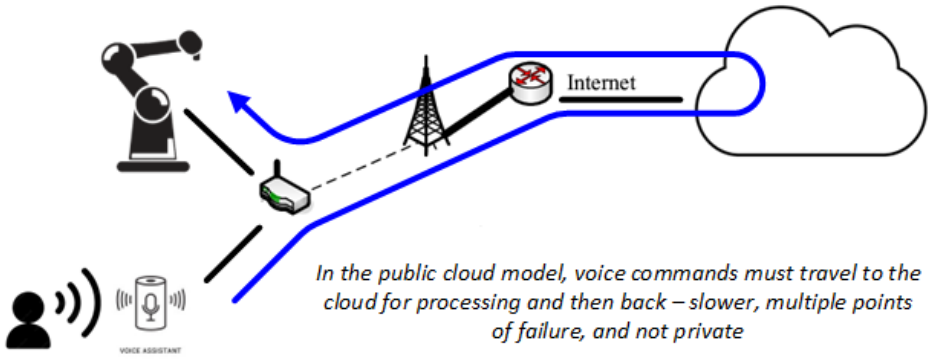
- **Failure prediction**
  - 24/7 failure prediction is required for robot maintenance and human safety
  - relies on vibration and acceleration data from 9-axis sensor data
  - also incorporates acoustic information (microphones)





Use Case Description, cont.

- **Speech recognition**
  - robots must recognize “immediate and urgent” voice commands in order to prioritize human safety
  - if someone shouts “stop now” the robot must stop — regardless of who is the speaker, level of background noise, or other circumstance
  - robots must accept verbal instructions. As examples, a factory with a rugged, wet, and fast-paced environment cannot use keyboards and phone apps; a robotaxi must react to alerts shouted by nearby drivers
- **No cloud dependencies**
  - cloud has many points of possible failure
  - we must be able to instruct robots and automated cars in critical situations, *regardless of cloud connectivity*

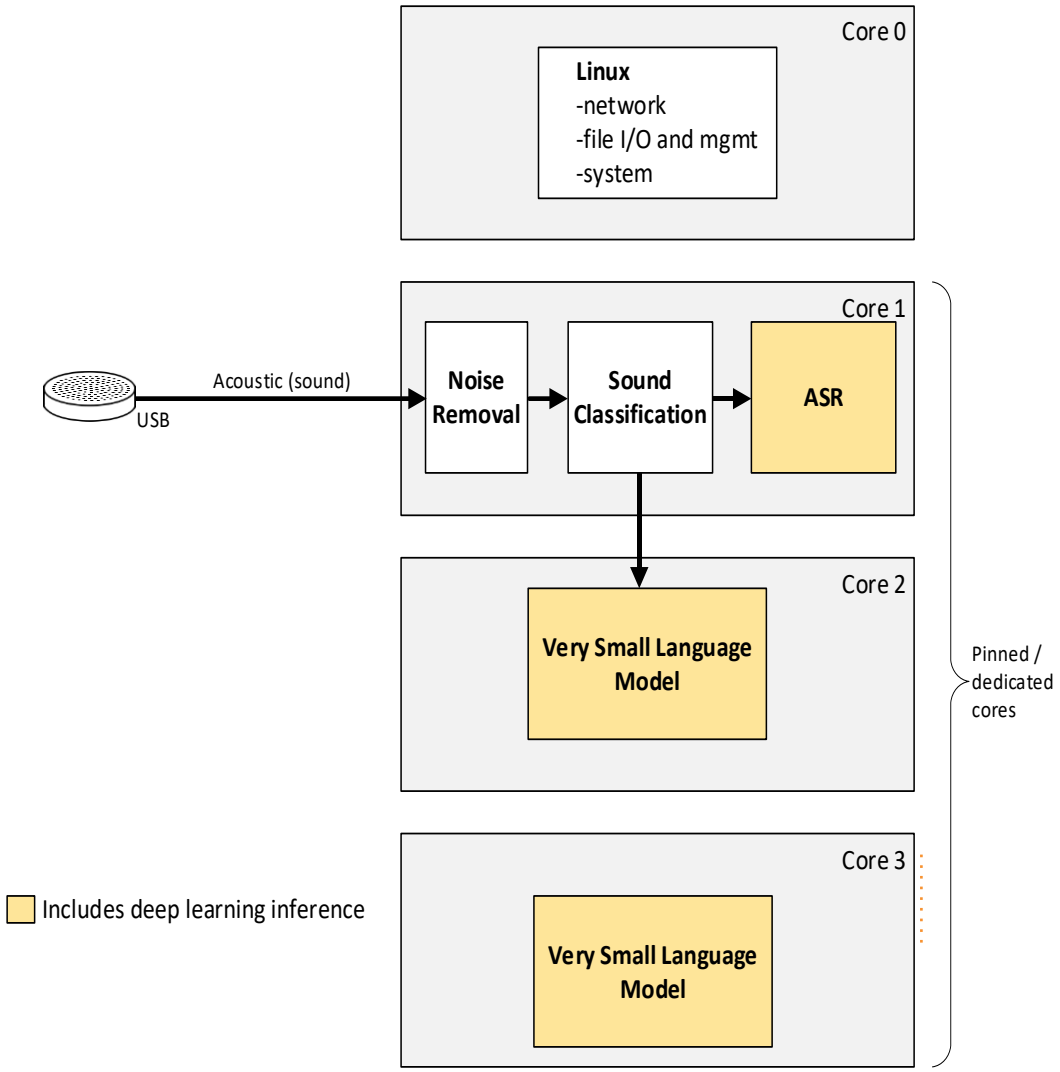


In the public cloud model, voice commands must travel to the cloud for processing and then back – slower, multiple points of failure, and not private

In the onboard HPC model, voice commands are processed by the robot

Current Work Area

- **Very Small Language Model (VSLM)**
  - robots must recognize conversational commands, without restrictions on vocabulary and word sequence
  - the typical speech recognition engine (e.g. Kaldi, Whisper) tends to generate frequent sound-alike errors. For example “a King ruled the state” might be recognized as “a King rolled the stake”. ChatGPT can easily fix such errors, but we need a very small model, not 1000s of GPUs
  - we have two (2) x86 cores and 8 GB mem available for this
- **Which model ?**
  - currently no small model that we tested can do this
  - we are investigating further and also working internally on alternative types of training that may reduce the model size



Achievements & Results

- **Multiple sensor prototype**
  - handles dishes without breakage
  - 9-axis sensor electronic design
  - real-time ARM + FPGA SoC processor
  - progress ongoing in reducing size, weight, and energy use

- **Roomba demo**
  - voice commands to control a Roomba
  - noise removal (vacuum, servo motors)
  - quad-core pico ITX server with no fans, “dead bugged” on a Roomba 680
  - no cloud dependency, no separate voice assistant (i.e. no Alex, Cortana, etc.)

About Akraino

- **LF Edge project**
  - open source community creating application blueprints for Edge Computing
  - spanning edge computing use cases in 5G, AI, Edge IaaS/PaaS, IoT, and robotics
  - all blueprints tested by the community and ready for adoption as-is, or can be used starting points for customizing new blueprints
- **To get involved**
  - Akraino – contact the Technical Steering Committee at tsc@lists.akraino.org
  - CPS Blueprint Family – contact fukano.haruhisa@fujitsu.com or jbrower@signalogic.com