

AI in Embedded Systems

Class 4: Systems

February 27, 2020

Louis W. Giokas

This Week's Agenda

Monday	Overview and Requirements
Tuesday	Toolkit - Algorithms
Wednesday	Toolkit - Hardware
Thursday	Systems
Friday	Application Examples

Course Description

This topic will cover a new approach to developing embedded systems which includes AI/Machine Learning(ML) as an element of the suite of tools available. As microcontroller devices become more powerful, these techniques are now within reach. We will look at the types of AI/ML algorithms available and appropriate in embedded systems. We will also look at how interaction with higher level systems, such as cloud analytics, can be integrated to create systems that evolve and improve over time and space.

Today's Agenda

- Overview
- Requirements
- Selecting Algorithms
- Selecting Devices
- System Architecture
- Conclusion/Next Class

Overview

- Embedded systems are just that, systems that interact with the environment and other systems in a complex way. This includes cloud computing resources, edge processing and the communications to tie it all together. We will look at new devices and capabilities in this area.

Overview

- Using a systems engineering approach is critical to developing an embedded AI application
- There are many moving parts
 - Algorithms
 - Hardware
 - Communications
- All must be selected in an ordered way to achieve the system goal

Requirements

- The starting point is the set of requirements for the system
 - Goals
 - External operating environment
- Drives the selection of algorithms which drives the selection of hardware
- Ensuring that these are understood will lead to a successful project

Requirements

- Understanding the external environment is critical to the selection process
 - Constrains the types of sensors that can be used
 - Sensors are how we interact with the environment, often a key part of embedded systems
 - Imposes constraints on size and power consumption of devices available
 - Many applications get around this by using communications to access additional compute resources

Requirements

- Always try to select the most efficient (smallest, lowest power, etc.) device for the particular application
- Select the algorithm that requires the least computational resources
- Match the communications to the task at hand
 - Don't require networking if the project can work stand-alone

Selecting Algorithms

- Different tasks call for different algorithms
 - Simple classification and prediction tasks can use ML statistical methods such as Least Squares, Linear Regression, Linear Discriminant Analysis (LDA), Logistic Regression, Tree Based Methods
 - Techniques such as Deep Learning may be required when processing image data
 - Parameter updates may be needed during operation
 - Can be done locally or globally

Selecting Algorithms

- A critical task is deciding how to handle training
 - One-time vs continuous
 - On-board vs remote
 - Remote may entail merging data from multiple systems or sensors
 - On-board will require enough memory to keep enough history
 - Storage of parameters

Selecting Algorithms

Simple Anomaly Detection Examples

$$\chi^2 = \sum_{i=1}^n \frac{(X_i - E_i)^2}{E_i}$$

The Chi-squared statistic is computed, and a large value denotes that the observed sample contains anomalies. The E_i are obtained from the training data. This system pattern is analogous to many IoT application Areas.

$$z = \frac{|\mathbf{x} - \bar{\mathbf{x}}|}{s}$$

Test score z computed for each new point

$$z > \frac{N-1}{\sqrt{N}} \sqrt{\frac{t_{\alpha/(2N), N-2}^2}{N-2 + t_{\alpha/(2N), N-2}^2}}$$

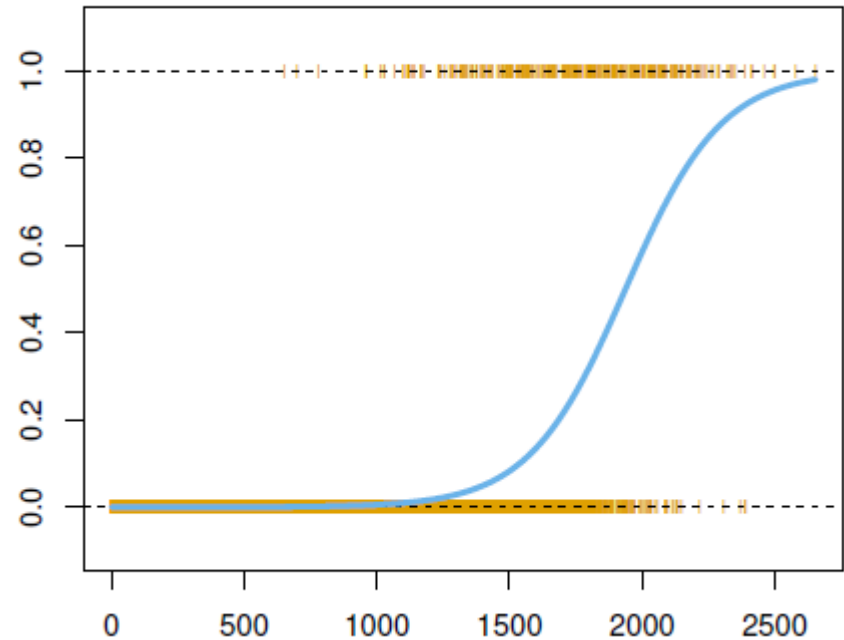
Compute the right hand side “once” and compare to z to see if this is an outlier. Can be done real-time at the edge.

Selecting Algorithms

$$p(X) = \Pr(Y = 1|X)$$

$$p(X) = \frac{e^{\beta_0 + \beta_1 X}}{1 + e^{\beta_0 + \beta_1 X}}$$

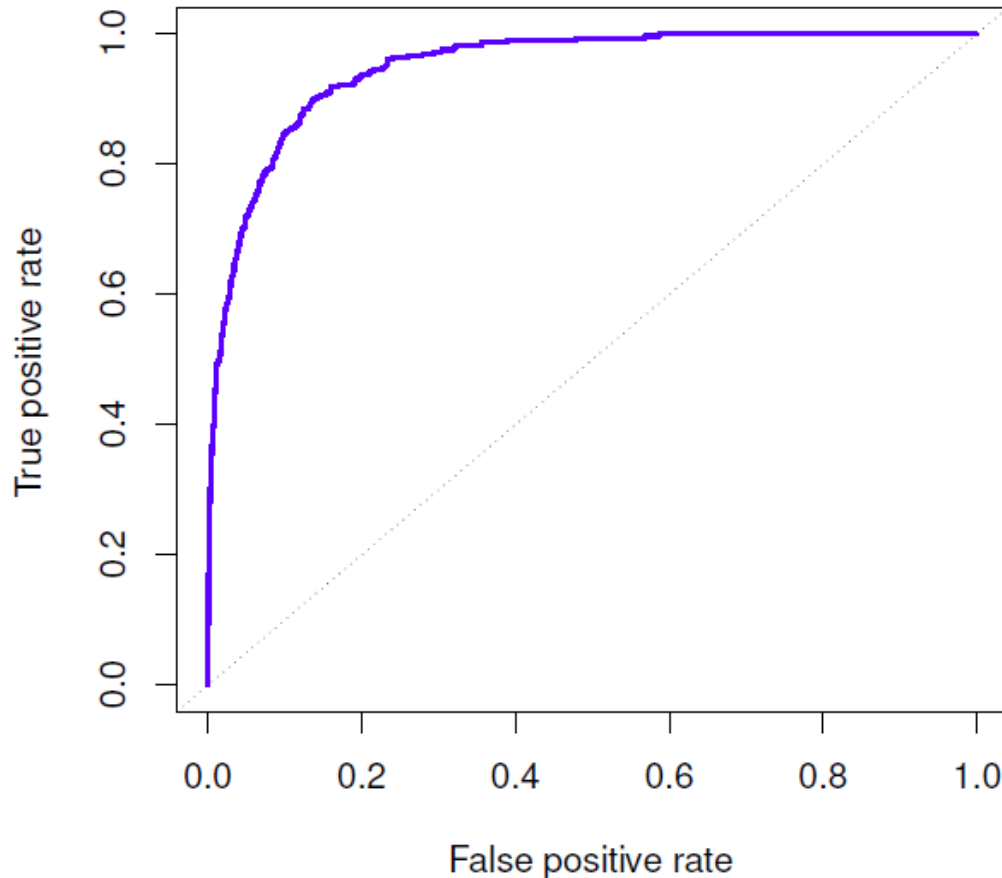
Logistic Regression used for classification



$$\log \left(\frac{p(X)}{1 - p(X)} \right) = \beta_0 + \beta_1 X.$$

Selecting Algorithms

ROC Curve

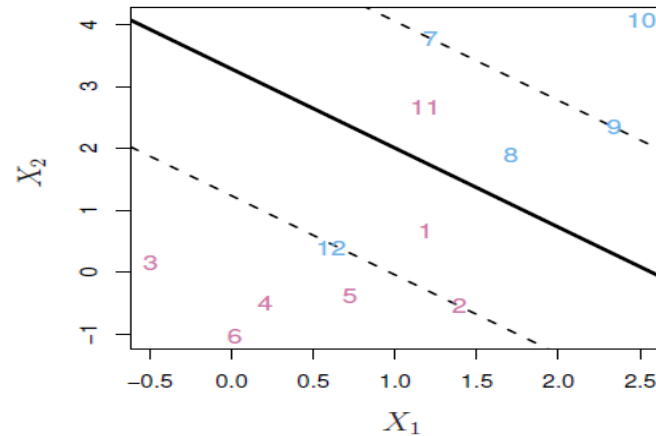
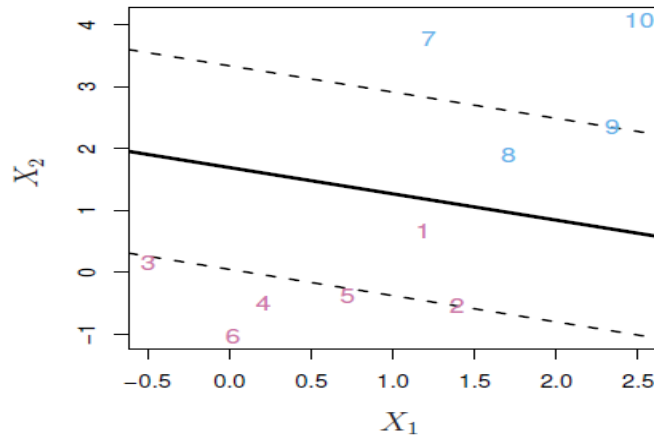


Linear Discriminate Analysis (LDA)

We can test the goodness of our model by the ROC curve and adjust parameters as necessary

Selecting Algorithms

Support Vector Machine (SVM) for classification



$$\text{maximize}_{\beta_0, \beta_1, \dots, \beta_p, \epsilon_1, \dots, \epsilon_n} M \quad \text{subject to} \quad \sum_{j=1}^r \beta_j^2 = 1,$$

$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}) \geq M(1 - \epsilon_i),$$

$$\epsilon_i \geq 0, \quad \sum_{i=1}^n \epsilon_i \leq C,$$

Presented by:

Selecting Devices

- Once the required algorithm has been selected, one must map this to the proper hardware platform
- There are a number of options available from single chips to board level systems
- Packaging is critical in embedded systems
 - Be aware of the environmental as well as processing constraints

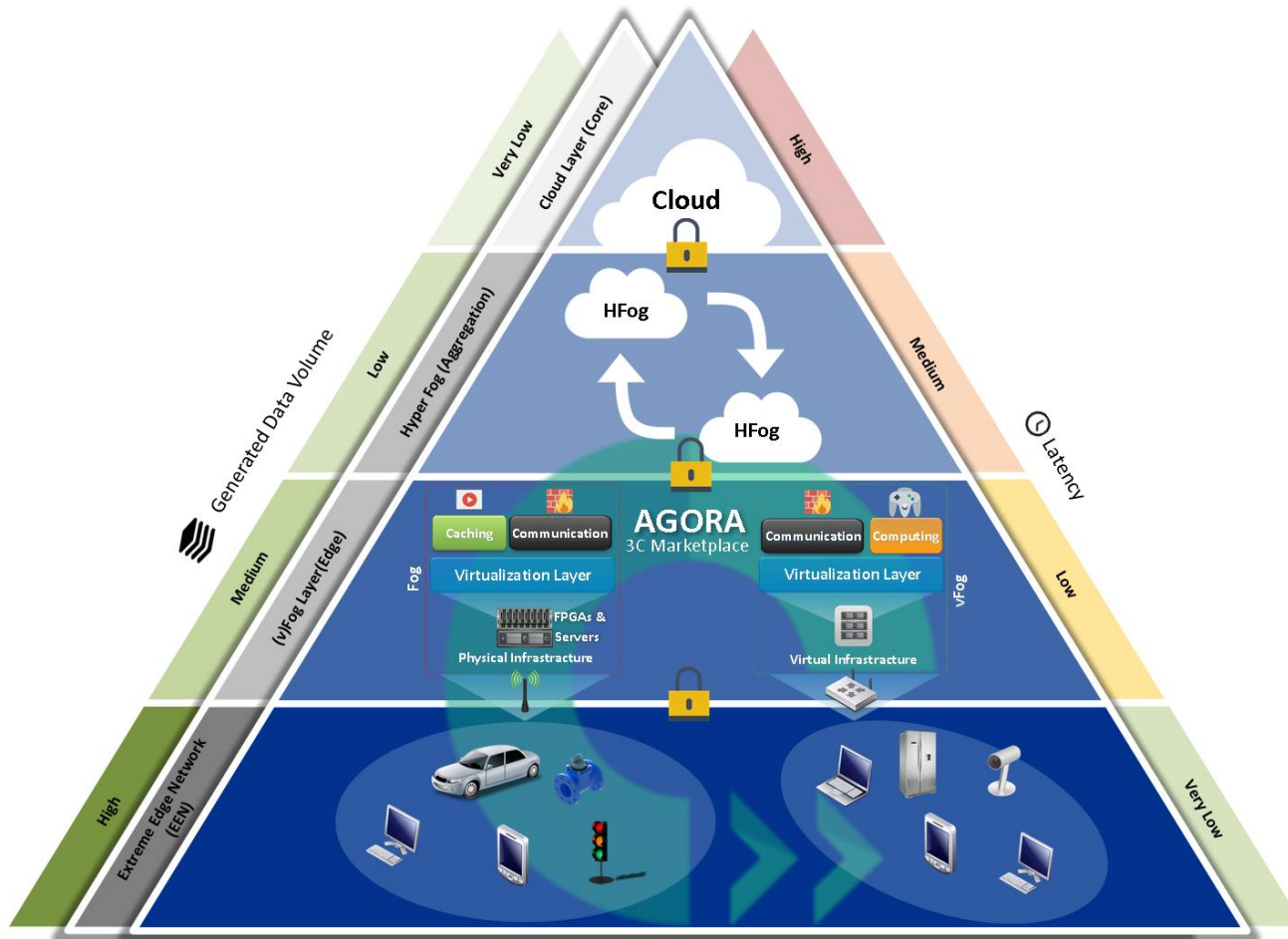
Selecting Devices

- Many vendors provide embedded devices with ML/DL capabilities
 - MicroSemi
 - STMicro
 - Cypress
 - NXP
 - Silicon Labs
- Choice of vendor will also depend on software support of the platform and ability to implement the requisite algorithms

System Architecture

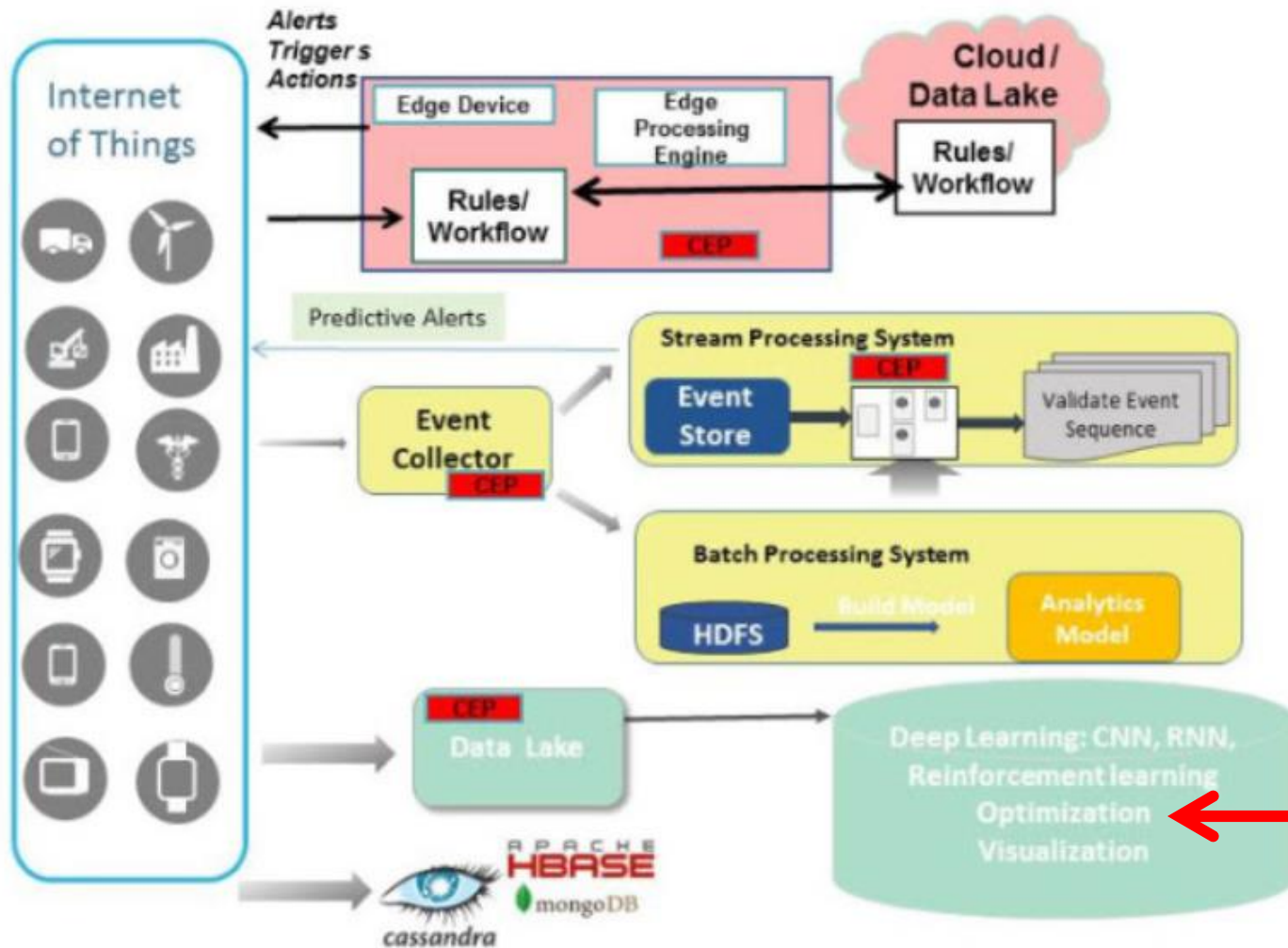
- As we have seen, architectures may be stand-alone or networked
- Stand-alone is the simplest and requires only a way to program and update
- Networked architectures may be single-tiered or multi-tiered
- Edge processing can be used in networked environments
- Cloud processing gives access to more data

Architectural Patterns



Presented by:

System Architecture

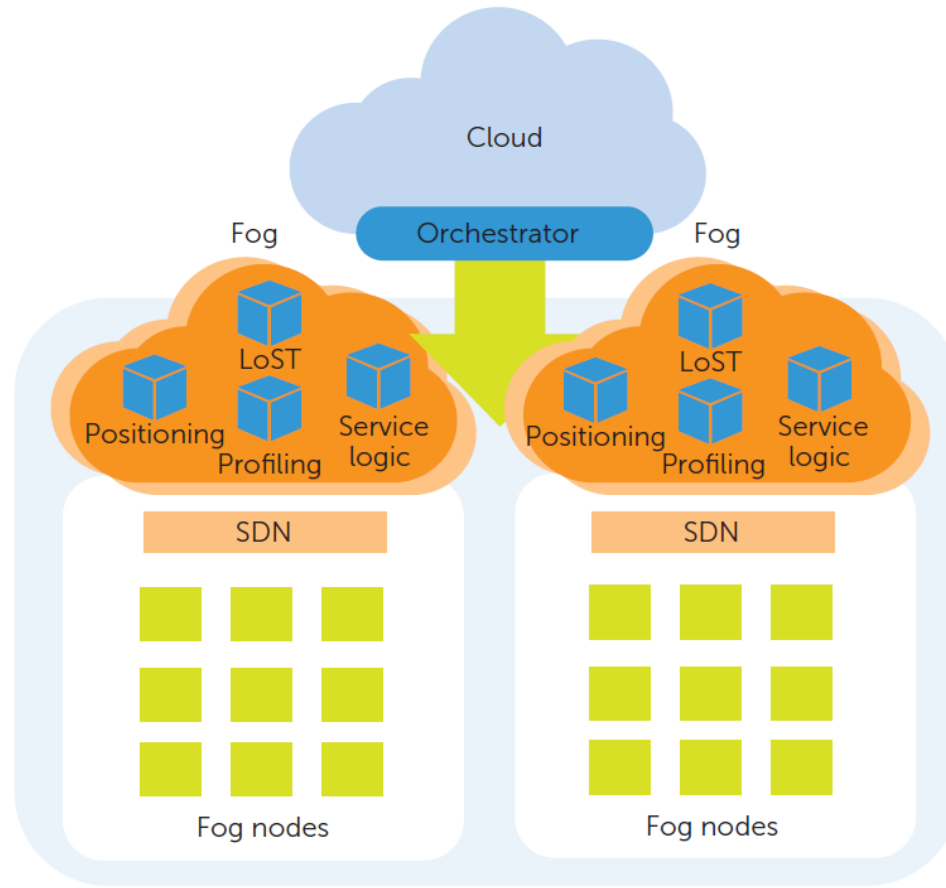


Optimization is a part of our IoT portfolio

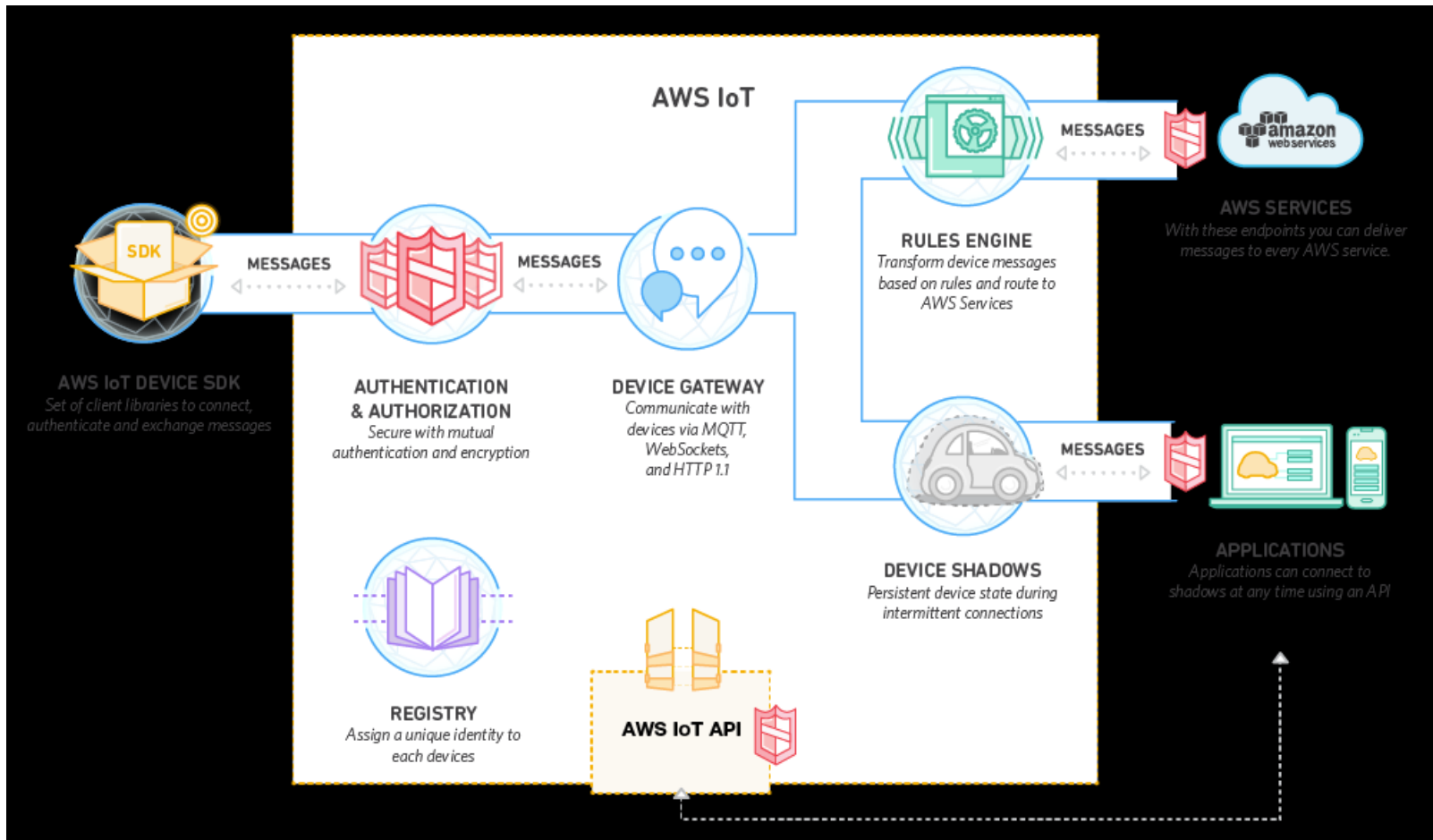
System Architecture



System Architecture



System Architecture



Presented by:

Conclusion/Next Class

- Today we have looked at the system aspects of AI in embedded systems
- We considered requirements, algorithm and device selection
- We reviewed some architectural considerations
- Tomorrow we will look at application examples