Internet of Things
IoT, Big Data and Other BS

Aslak Johansen asjo@mmmi.sdu.dk

Feb 4, 2025

SDU

Part 1:
Who We Are, and
Where We Go from Here

Meet the Students

You are 72 MSc students of Software Engineering?

You are expected to have the equivalent of a bachelor in Software Engineering.

We will touch on

- Computer Systems (bit, byte, bit shifting, memory, bus)
- Concurrency (threads, race conditions)
- Operating Systems (process management)
- Statistics (basic stuff)
- Electronics (basic stuff)
- Electromagnetism (superficially)

Survey

https://PollEv.com/surveys/TDR7aQ31EDCV7XHolGP5P/respond

Meet the Course

IoT Systems, covering the following areas (more or less):

- Data collection: Sensing on device level
- Data transportation
- Information (or metadata) management
- Data processing
- Decision support
- Realizing a decision (actuation)

You are going to work on these areas in the context of IoT.

Meet the Teacher: Aslak Johansen

Background in experimental computer science

PhD in energy-harvesting wireless sensor networks

Currently working on

- Concurrency, availability and predictable responsiveness in systems
- Stream and in-network processing
- Demand-driven data processing
- (Meta)data /information modeling
- Environmental sensing
- Interaction models
- Building operating systems

Despite this, I am not always boring at parties.

Supervised MSc Projects (shameless plug, I know)
https://github.com/aslakjohansen/student-projects

Intended Learning Outcomes

Knowledge: The student

- is able to describe the fundamental terminology of IoT systems.
- is able to list and describe the dimensions of the wireless performance space.

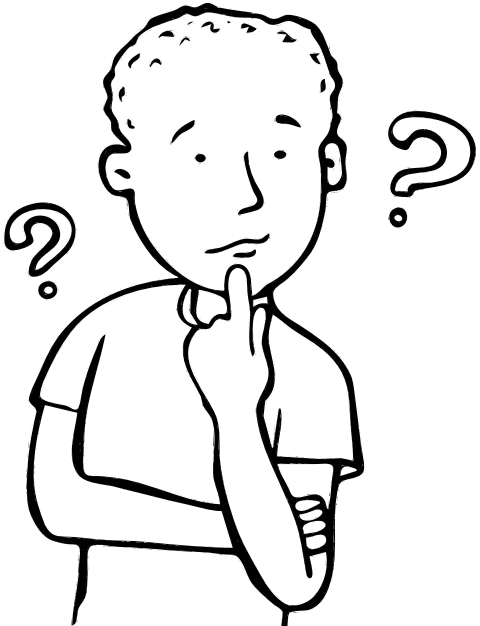
Skills: The student

- is able the describe the architecture of IoT systems.
- is able to design and implement IoT systems.
- is able to measure and present relevant performance characteristics.

Competences: The student

- is able to analyze data collection requirements and map them to a data collection strategy covering sampling and budget management.
- is able to compare and argue for the choice of IoT-relevant radio networking technologies.
- is able to reflect of the privacy and resource implications of the placement of logic within an IoT system.

Questions?



https://openclipart.org/detail/238087/boy-thinking-of-question

Part 2:
The Big Picture

Internet of Things

What is the Internet of Things?

Suggestions?

Internet of Things ▷ Definition

a buzzword!

Definitions often involve:

- Gadgets
- A range of everyday items which have been networked and are controllable from a single app
- Everyday items which have been instrumented
- *"Things"* that have been networked

What is a *"thing"* and what does it mean to be *"networked"*?

In this course we define IoT as:

"A cyber-physical system consisting of a large number of individually and globally addressable devices which have a significant portion of their logic running elsewhere"

Internet of Things ▷ Consequences

Long list of players, each with their own agenda and varying background

This leads to ...

1. siloed access (*"I want the whole cake!"*)
2. lots of external dependencies (*"I have to integrate with that as well?"*)
3. lack of security (*"first to market!"*)

In this course we

- address the first two
- skip the last one
  - Why?

Big Data

What is Big Data?

Suggestions?

Big Data ▷ Definition

a buzzword!

Definitions often involve:

- *"Logic dealing with a large amount of data"*
- Requirements what cannot (efficiently) be served by a single kernel, due to at least one of the following:
  - Processing requirements
  - Memory requirements
  - Network requirements
- Limitations in the capacity of traditional processing methods (e.g., relational DBMS)
- Complex logic for combining a large amount of data sources

In this course we define big data as:

"A class of processing regimes which primary property is that it relies on a body of data which expands at a rate where processing has become nontrivial"

Big Data ▷ Consequences

Often, the idea behind big data is that all kinds of data are collected with the hope that patterns will eventually emerge and be exploitable.

This leads to

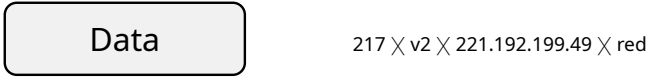
- more sampling hardware
- higher network capacity use
- larger databases
- more processing

This, in turn, affects

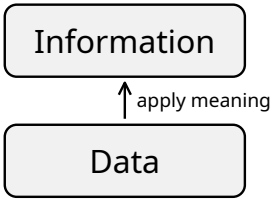
- system design
- energy consumption
- bill of materials
- price

Enriching Data to Wisdom

Enriching Data to Wisdom

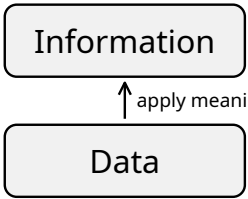


Enriching Data to Wisdom



217 ╳ v2 ╳ 221.192.199.49 ╳ red

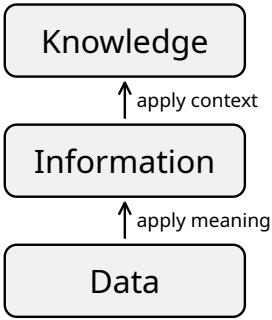
Enriching Data to Wisdom



North-facing light on Niels Bohrs Alle and Campusvej has turned red

217 ╳ v2 ╳ 221.192.199.49 ╳ red

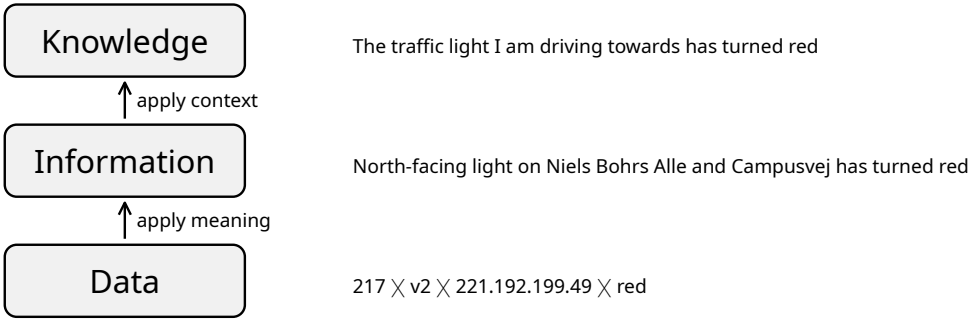
Enriching Data to Wisdom



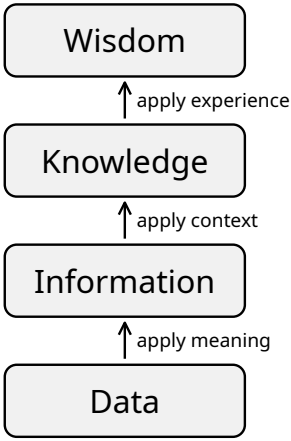
North-facing light on Niels Bohrs Alle and Campusvej has turned red

217 ╳ v2 ╳ 221.192.199.49 ╳ red

Enriching Data to Wisdom



Enriching Data to Wisdom

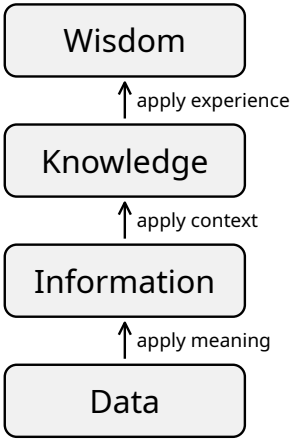


The traffic light I am driving towards has turned red

North-facing light on Niels Bohrs Alle and Campusvej has turned red

217 ╳ v2 ╳ 221.192.199.49 ╳ red

Enriching Data to Wisdom



I better apply the brakes!

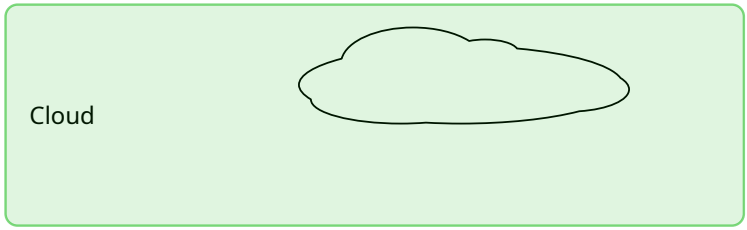
The traffic light I am driving towards has turned red

North-facing light on Niels Bohrs Alle and Campusvej has turned red

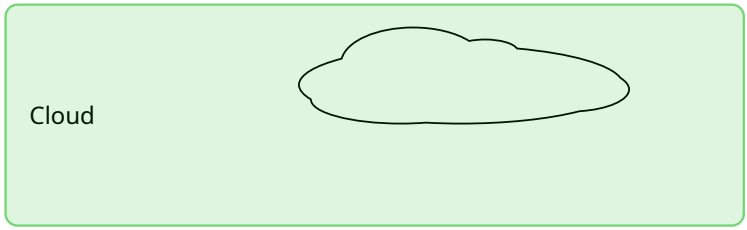
217 ╳ v2 ╳ 221.192.199.49 ╳ red

Cloud, Fog, Edge and Mist

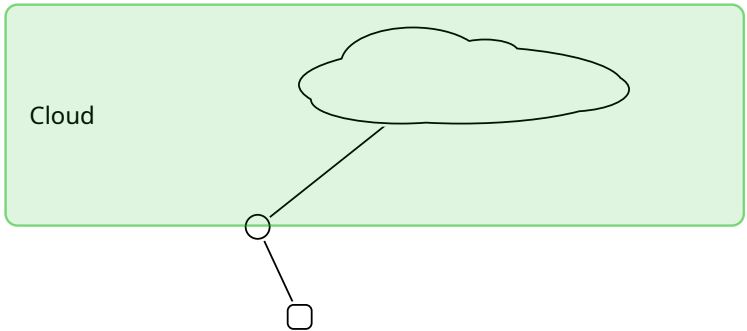
Cloud, Fog, Edge and Mist



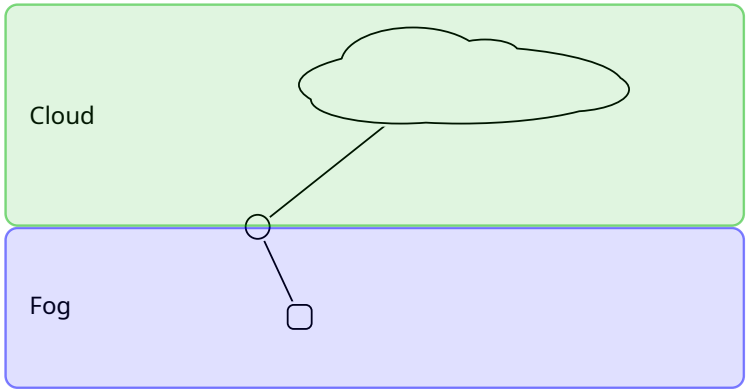
Cloud, Fog, Edge and Mist



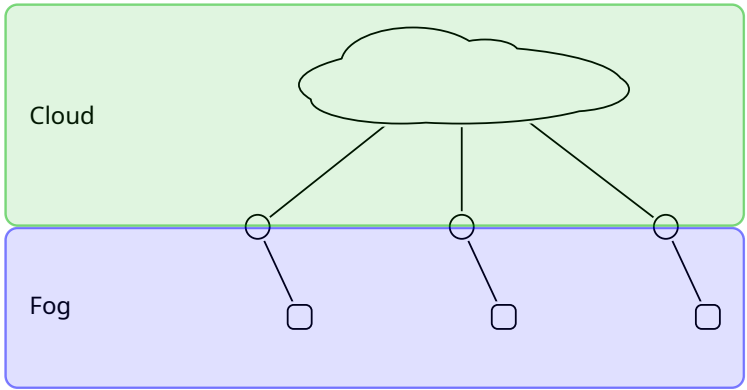
Cloud, Fog, Edge and Mist



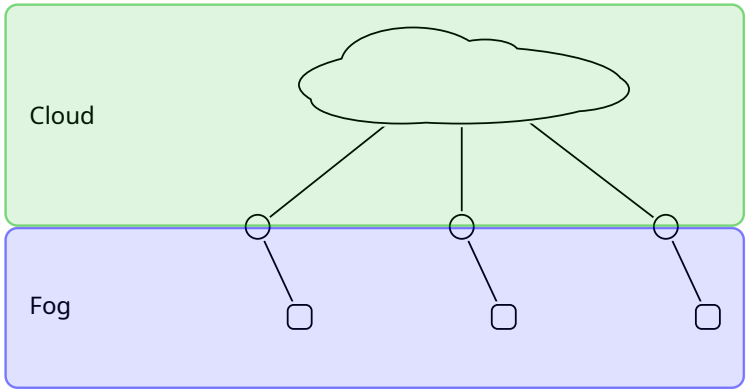
Cloud, Fog, Edge and Mist



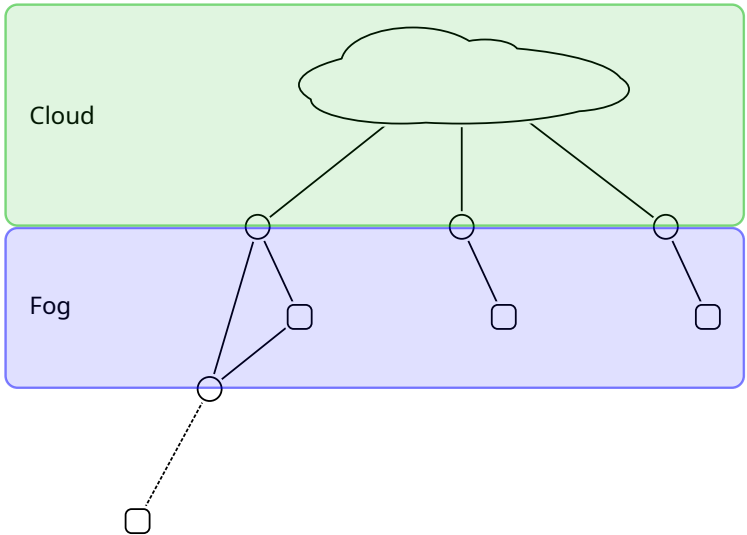
Cloud, Fog, Edge and Mist



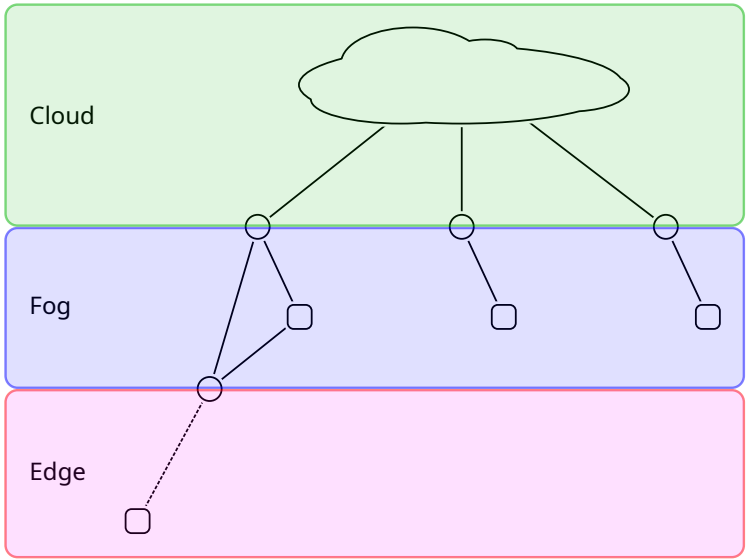
Cloud, Fog, Edge and Mist



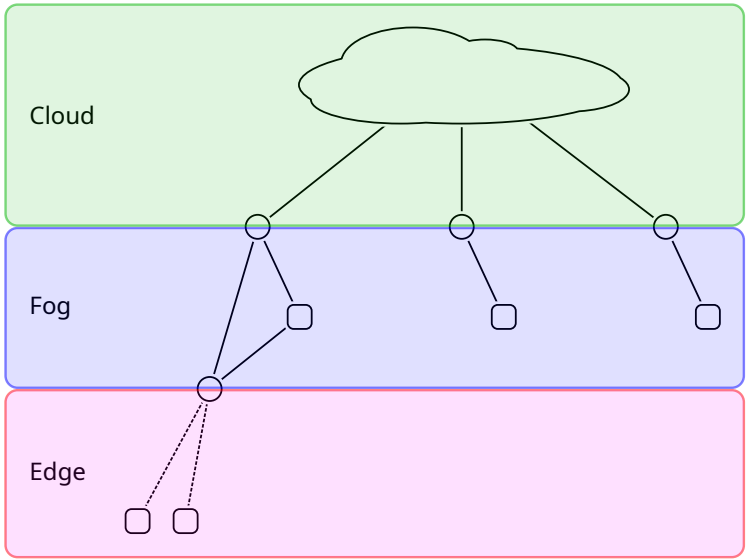
Cloud, Fog, Edge and Mist



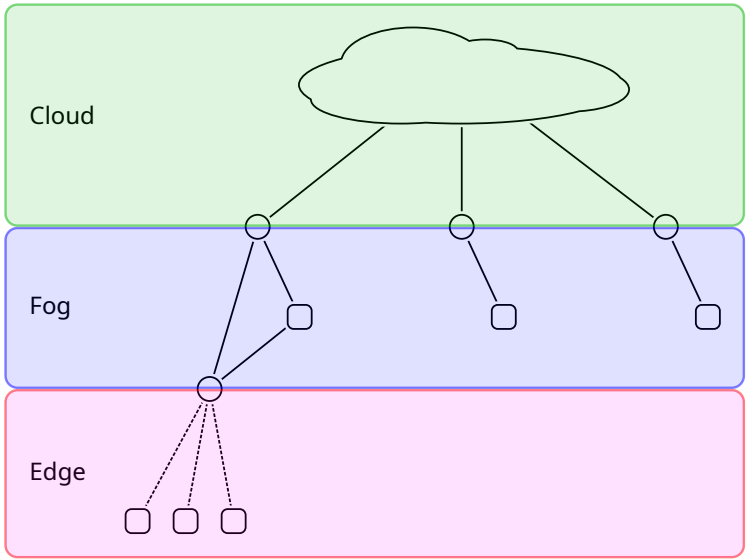
Cloud, Fog, Edge and Mist



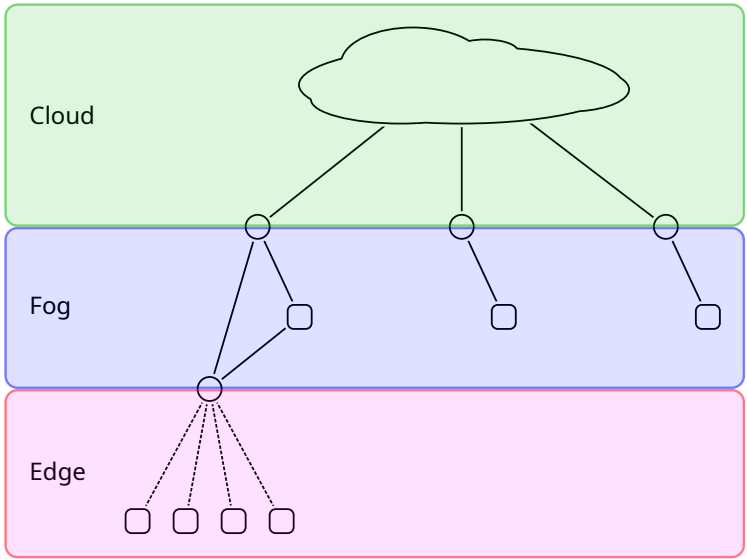
Cloud, Fog, Edge and Mist



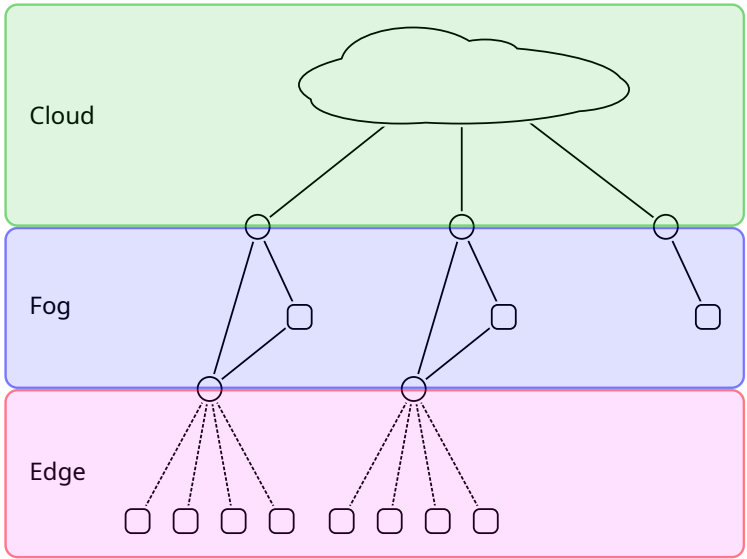
Cloud, Fog, Edge and Mist



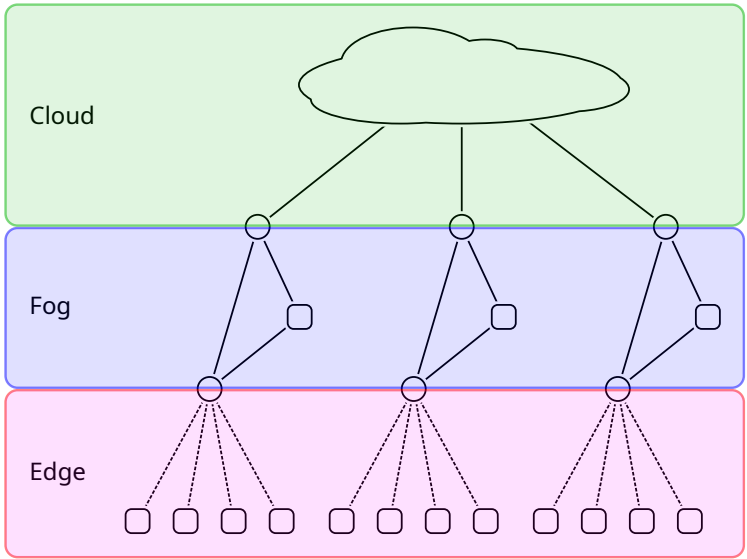
Cloud, Fog, Edge and Mist



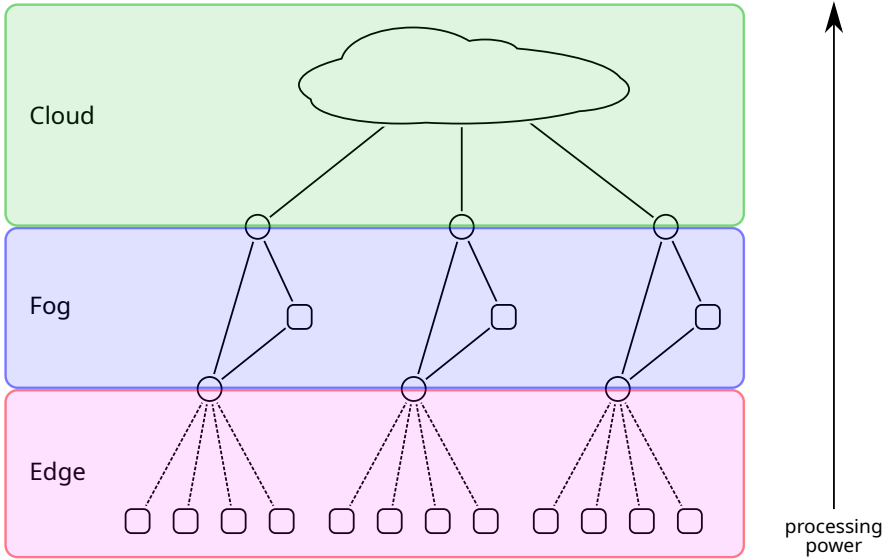
Cloud, Fog, Edge and Mist



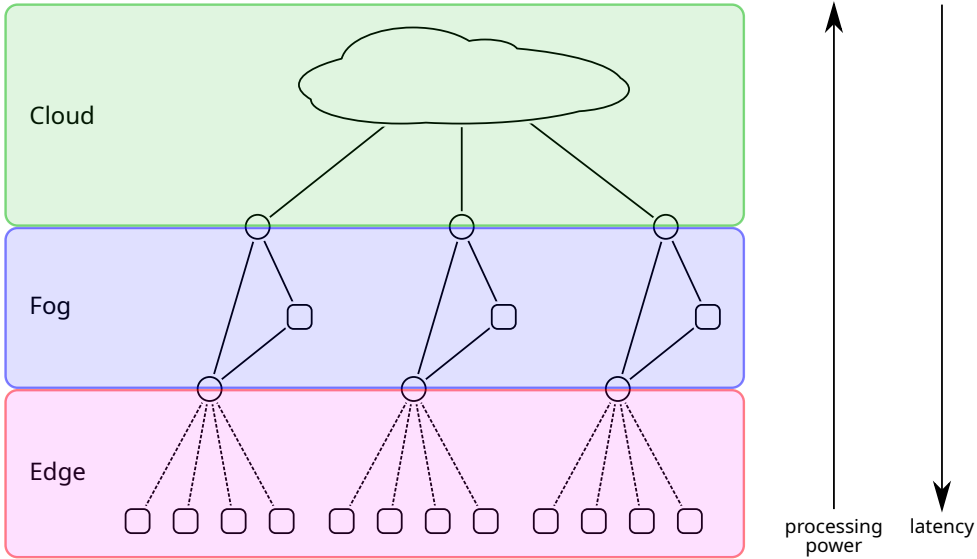
Cloud, Fog, Edge and Mist



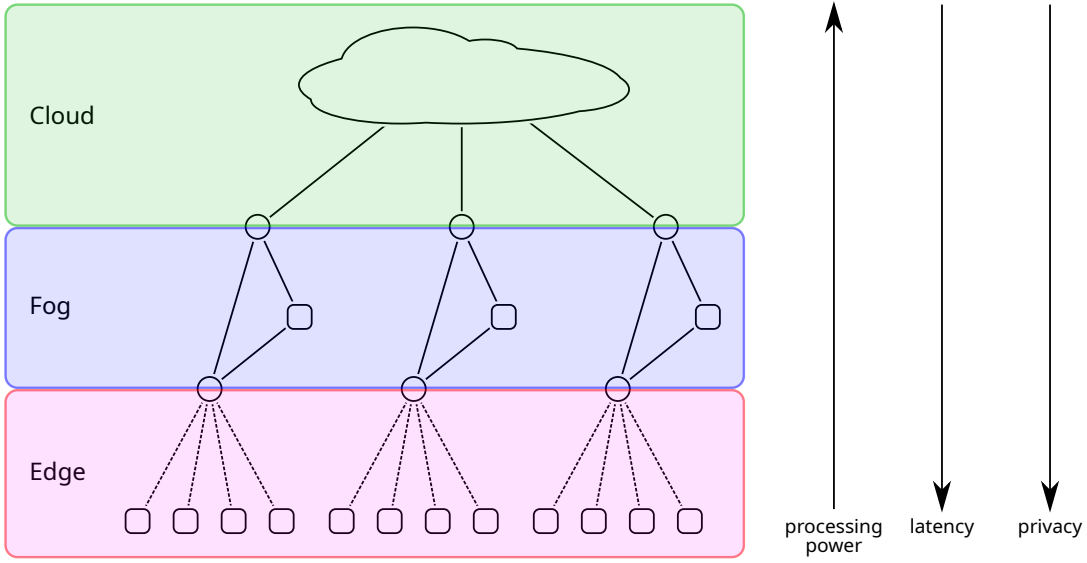
Cloud, Fog, Edge and Mist



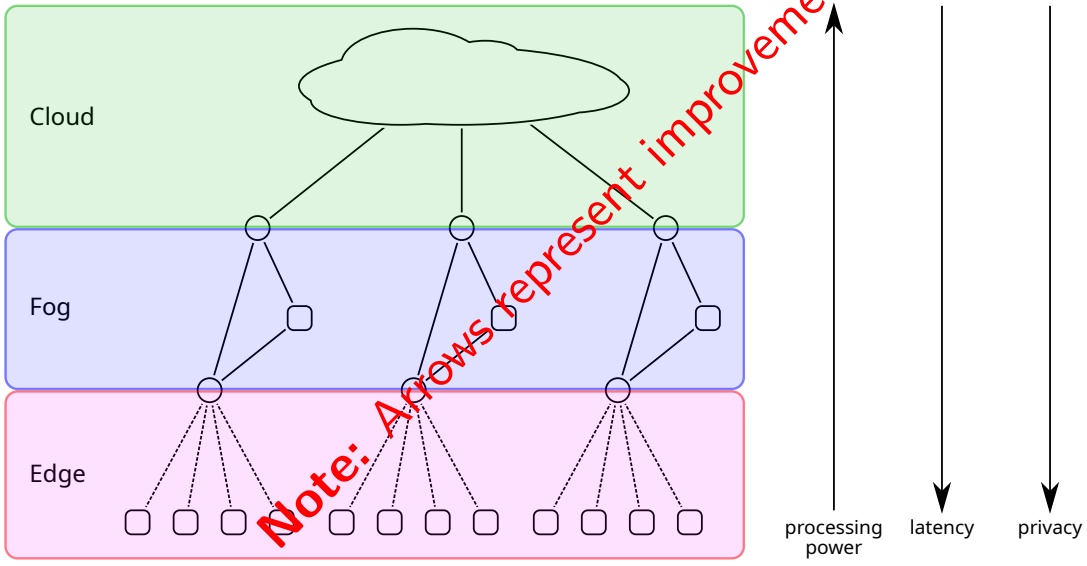
Cloud, Fog, Edge and Mist



Cloud, Fog, Edge and Mist



Cloud, Fog, Edge and Mist



A GDPR Primer

- ***Who*** EU citizens (similar rulesets have been introduced elsewhere).
- ***When*** Collected data which can in itself (or combined with other data) be linked to an individual, either directly or through processing.
- ***Consequences***
  - Individuals need to be informed of what data are being collected, for which purpose it is being collected and for which period of time it will be stored.
  - Individuals have to voluntarily agree to the collection.
    Implication: It is illegal to discriminate against individuals who have opted out.
- ***Violations*** Significant penalties.

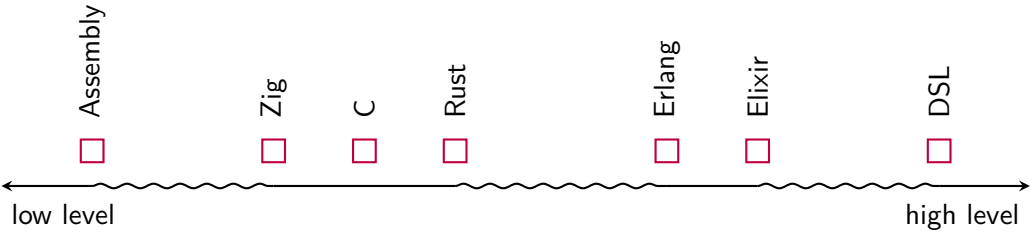
Programming Languages ▷ Edge

Requirements:

1. Device drivers require weak realtime properties ⇒ No stop-the world GC.
2. Depending on use-case, some amount of concurrency.

Requirement 1 excludes languages with a traditional (e.g., mark-sweep) garbage collector: Java, C#, Python, Go ...

Of the remaining ones, the following are relevant:



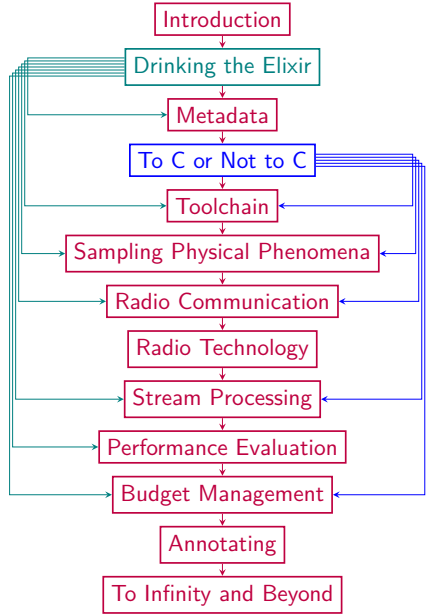
Programming Languages ▷ Cloud and Fog

Requirements:

1. Handle high amounts of concurrency.
   This excludes any language that builds its concurrency model on top of threads.
2. Have a high degree of availability.
   This excludes any language that does not support hot updates of code and granular supervision.
3. Offer robust low response times.
   This excludes any language that does not do high-granularity preemptive scheduling.
4. Be mature.
   This excludes languages without tooling to deal with these requirement during development and operation, and any language that hasn't been properly battle-tested. Notably, this excludes Unison.
5. Access to frameworks for modeling information (as graphs or through ontologies).

Of the remaining ones the two most popular BEAM-based languages stand out: Erlang and Elixir.

Course Overview



Welcome and overview

Elixir programming language

Describing configurations, deployments and datastreams

C programming language

Getting started with the ESP32 toolchain

Interfacing with sensors

Theory of radio communication and practical wifi

Other radio technologies and techniques

Working with streaming data

Evaluating non-functional requirements

Dealing with restrictions: cpu, memory, network, energy

State lifecycle: Identity, time and location of a device

Looking towards the future

Weekly Model

Tuesdays (4h):

- Preparations (generally) needed.
- Lecture introducing new material.
- Exercises with physical availability of lecturer.

Thursdays (4h):

- More time for same exercises.
- Physical space available: DevOps Lab (Ø41-510a-0) and Agile Lab (Ø40-606a-0).
- Lecturer available remotely (over discord).

Note 1: This leaves $\frac{280\text{h}}{13} - (4\text{h} + 4\text{h}) = 13.5\text{h}$ for preparation, additional work on exercises and exams.

Note 2: Exercises are important ... for your exams!

Exam Format

The exam consists of:

- Written
- MCQ

Note 1: The exam will refer to the exercises, and assume that you have personally gone through the involved thought processes.

Note 2: This does not mean that you are not allowed to work in groups, just that you are individually evaluated on all aspects.

Discussion: Intelligent vs Smart

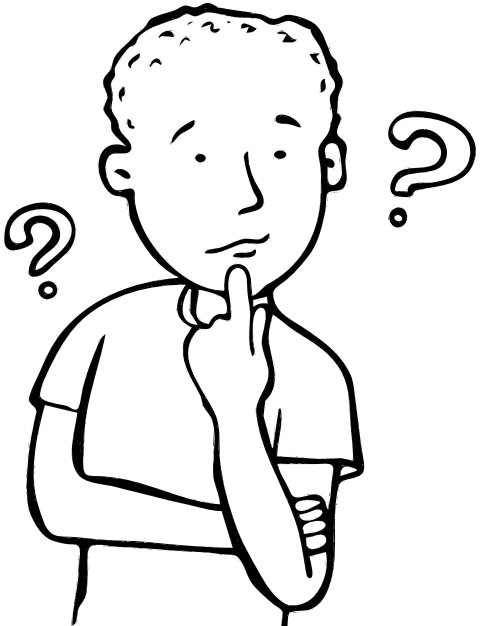
Typical IoT Products

- Smart Bulb
- Intelligent Coffee Can
- Smart Slow Cooker
- Intelligent Door Bell
- . . .

Open Questions:

- What does it mean for a device to be *"smart"*?
- Are any of the above mentioned devices *"smart"*?
- What does it mean for a device to be *"intelligent"*?
- Are any of the above mentioned devices *"intelligent"*?

Questions?



https://openclipart.org/detail/238087/boy-thinking-of-question