



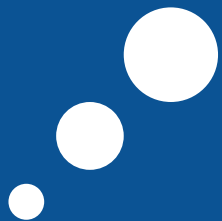
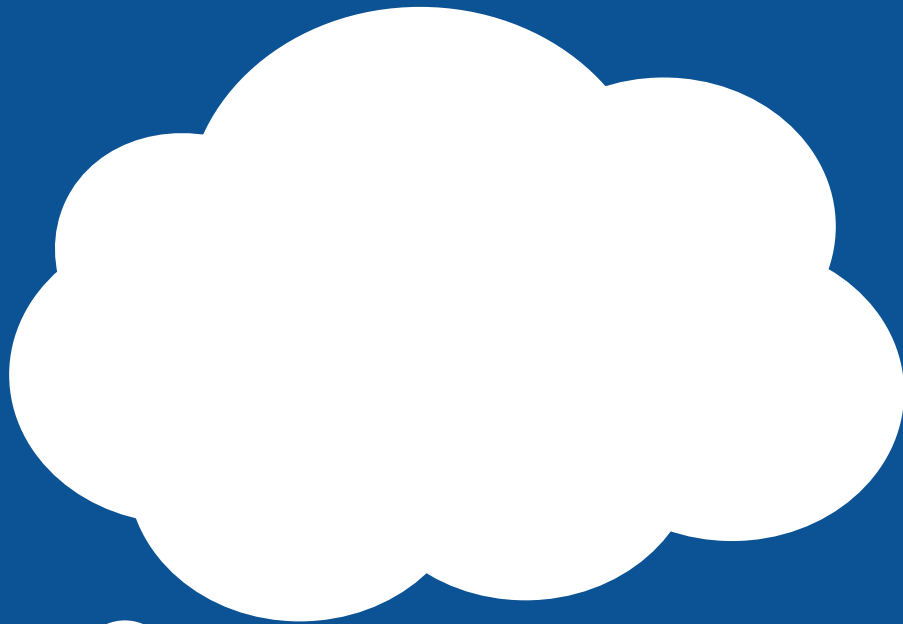
Living on the Cloud's Edge

Tannaz N. Roshandel

@tannaznvr

Rustam Mehmandarov

@rmehmandarov



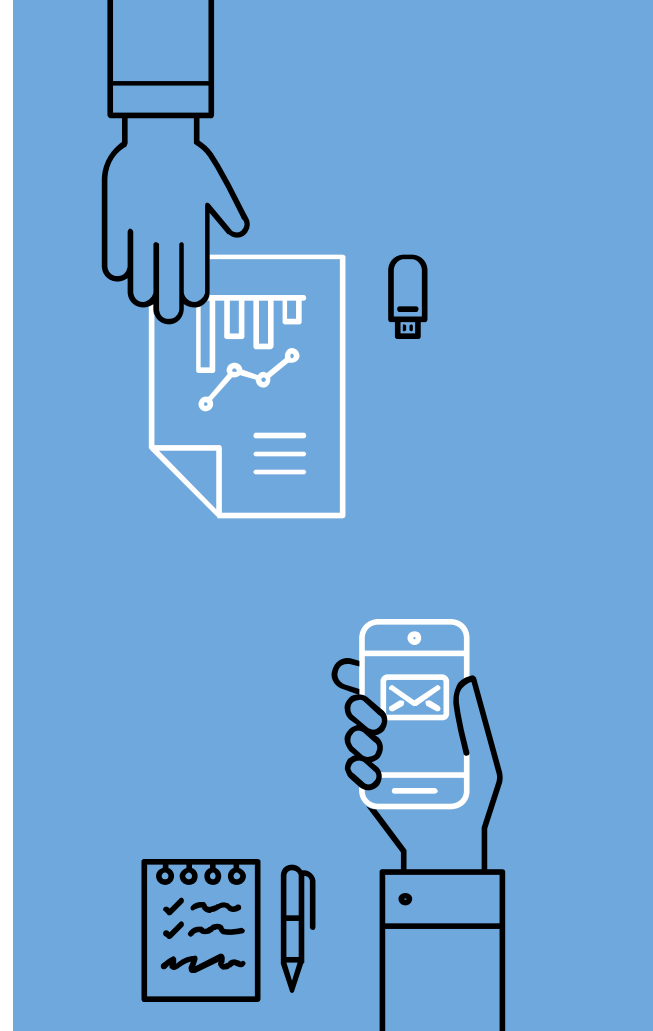
Building something...

 Comply with Strict Privacy Policy

 High Performance

 Quick (Low Latency)

 Cheap to Run



HELLO!

Rustam Mehmandarov

@rmehmandarov



computas

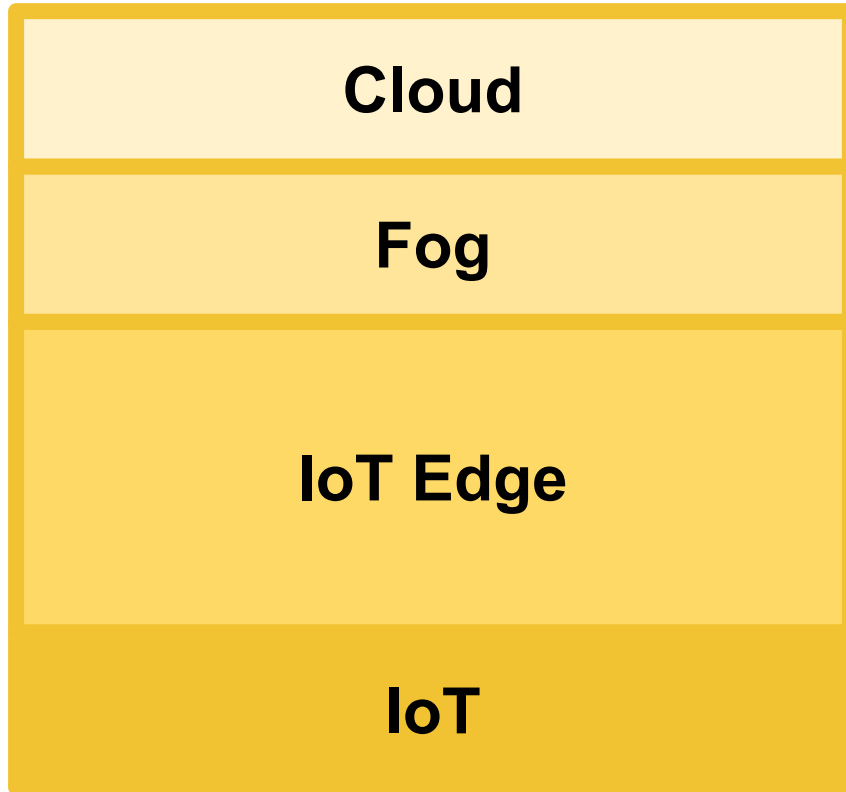
Tannaz N. Roshandel

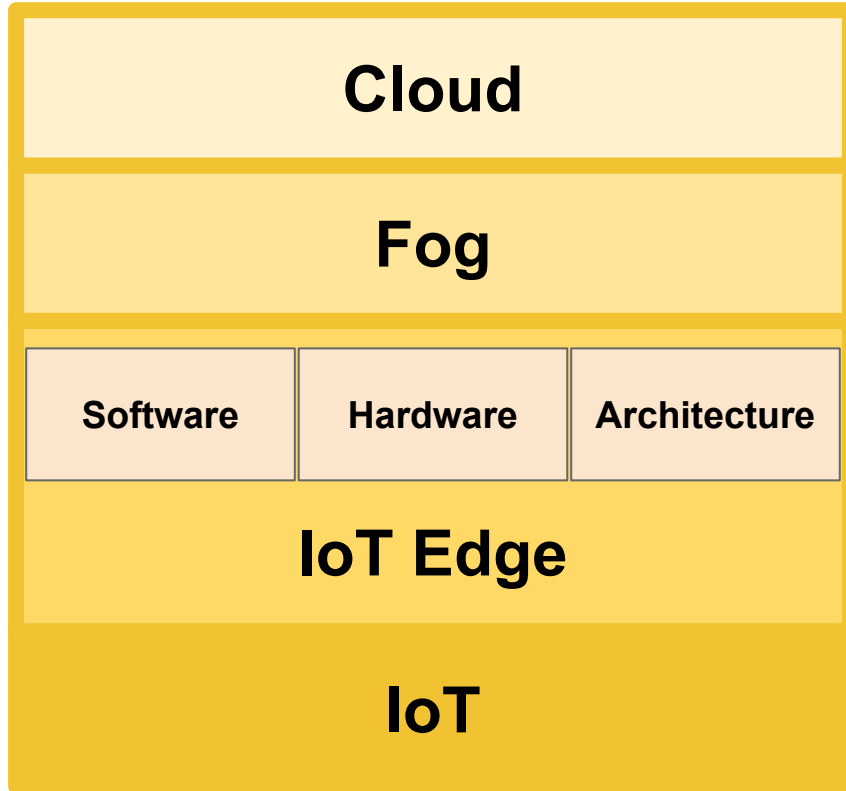
@tannaznvr

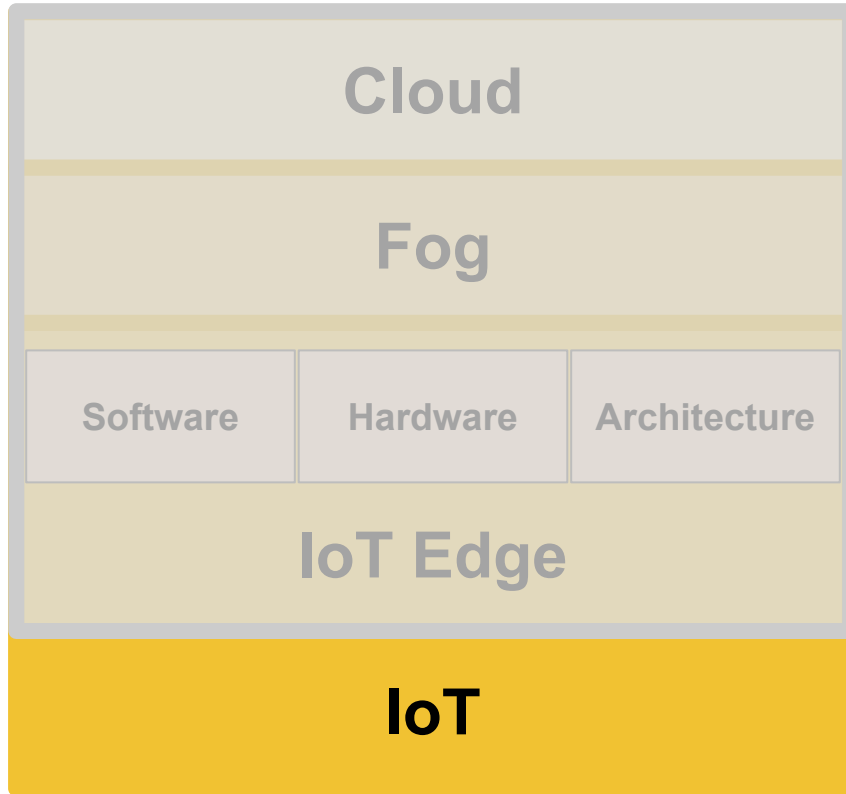


UiO • University of Oslo









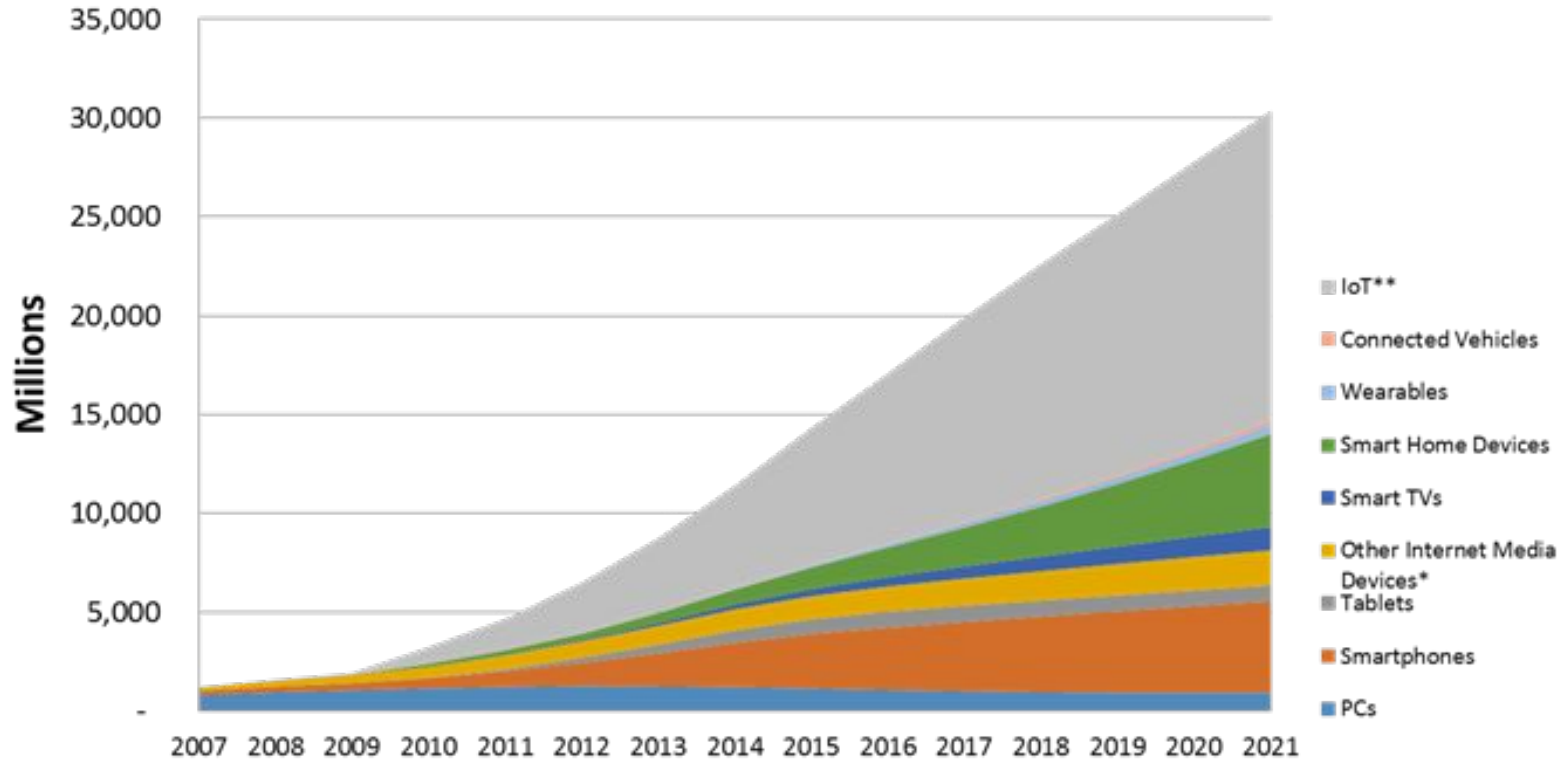
IoT

Internet of Things

Electronic devices with Internet connectivity that can be remotely monitored and controlled



Global Connected and IoT Device Installed Base Forecast



Source – Strategy Analytics research services ,October 2017: IoT Strategies , Connected Home Devices, Tablet and Touchscreen Strategies, Wireless Smartphone Strategies, Wearable Device Ecosystem, Smart Home Strategies

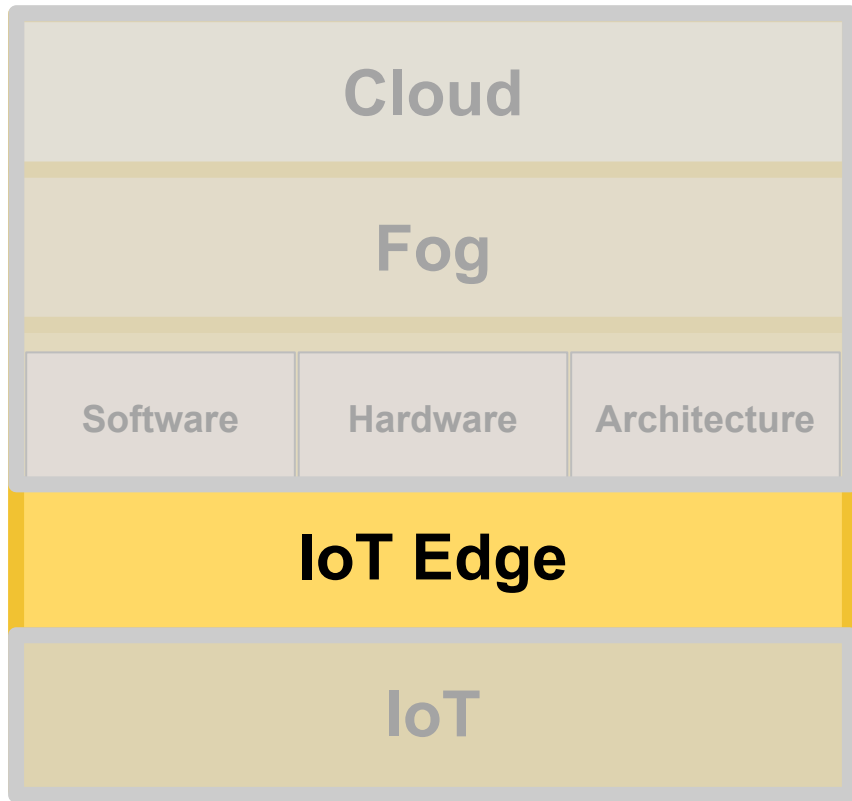
Internet of Things

+

- Lead time processing
- Accessibility
- Mobility
- Tracking Ability
- Efficiency
- Cost

-

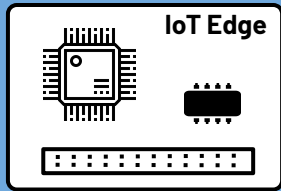
- Data Security
- Privacy
- Massivity of Data
- Environment Issues
- Complexity
- Network dependent



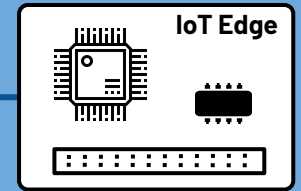
Edge Computing

Improve operations
and cut costs





Edge computing is a mesh network of **micro data centers** that **process or store** critical data locally and **push** it to a central data center or Cloud.



— IDC

International Data Corporation

Edge Computing

+

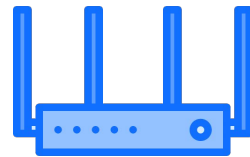
- Low Latency
- Privacy
- Real-Time Availability
- Real-Time Data Transmission
- Productivity Increase

-

- Limited (or no) Redundancy
- Potential Loss of Data
- Higher Risk in general



**IoT "Smart Devices", Assistants, etc.
(not Edge devices)**



Gateway



Cloud



Rich Rogers

@RichRogersloT

Følg



My wife asked me why I was speaking so softly at home.

I told her I was afraid Mark Zuckerberg was listening!

She laughed. I laughed.

Alexa laughed. Siri laughed.

17:00 - 1. jun. 2019

12 998 retweets 48 437 liker



346

13 K

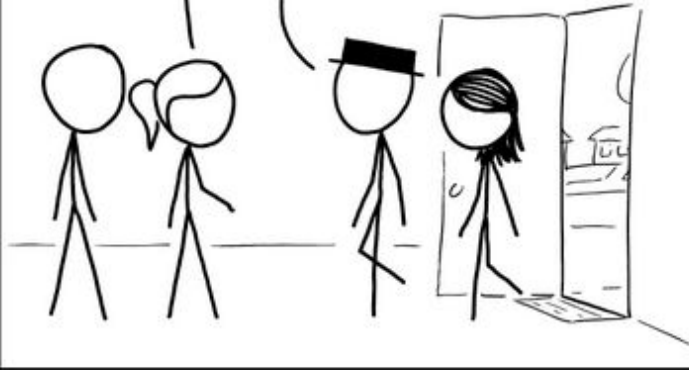
48 K

HELLO, WELCOME TO OUR HOUSE!

THANKS FOR INVITING US!

ALEXA, ORDER TWO
TONS OF CREAMED CORN.

ALEXA, CONFIRM PURCHASE.



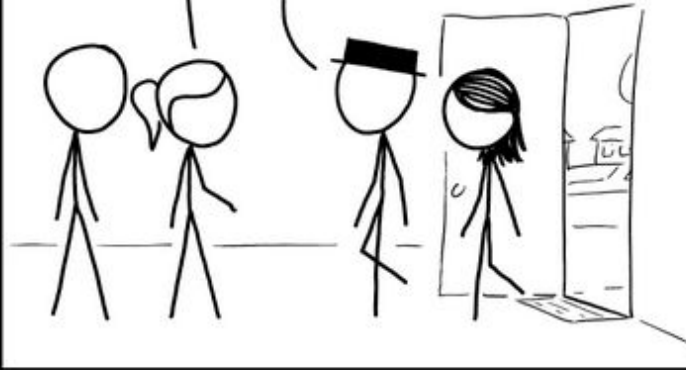
<https://xkcd.com/1807/>

HELLO, WELCOME TO OUR HOUSE!

THANKS FOR INVITING US!

ALEXA, ORDER TWO
TONS OF CREAMED CORN.

ALEXA, CONFIRM PURCHASE.

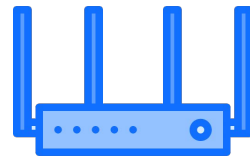


WHEN VISITING A NEW HOUSE, IT'S
GOOD TO CHECK WHETHER THEY HAVE
AN ALWAYS-ON DEVICE TRANSMITTING
YOUR CONVERSATIONS SOMEWHERE.

<https://xkcd.com/1807/>



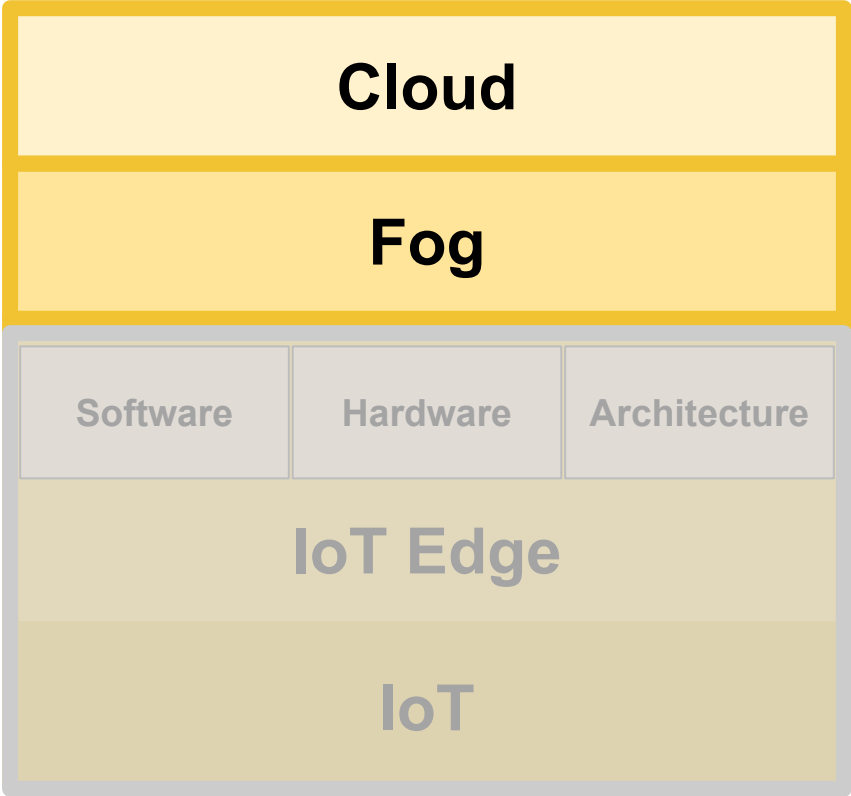
**IoT "Smart Devices", Assistants, etc.
(not Edge devices)**



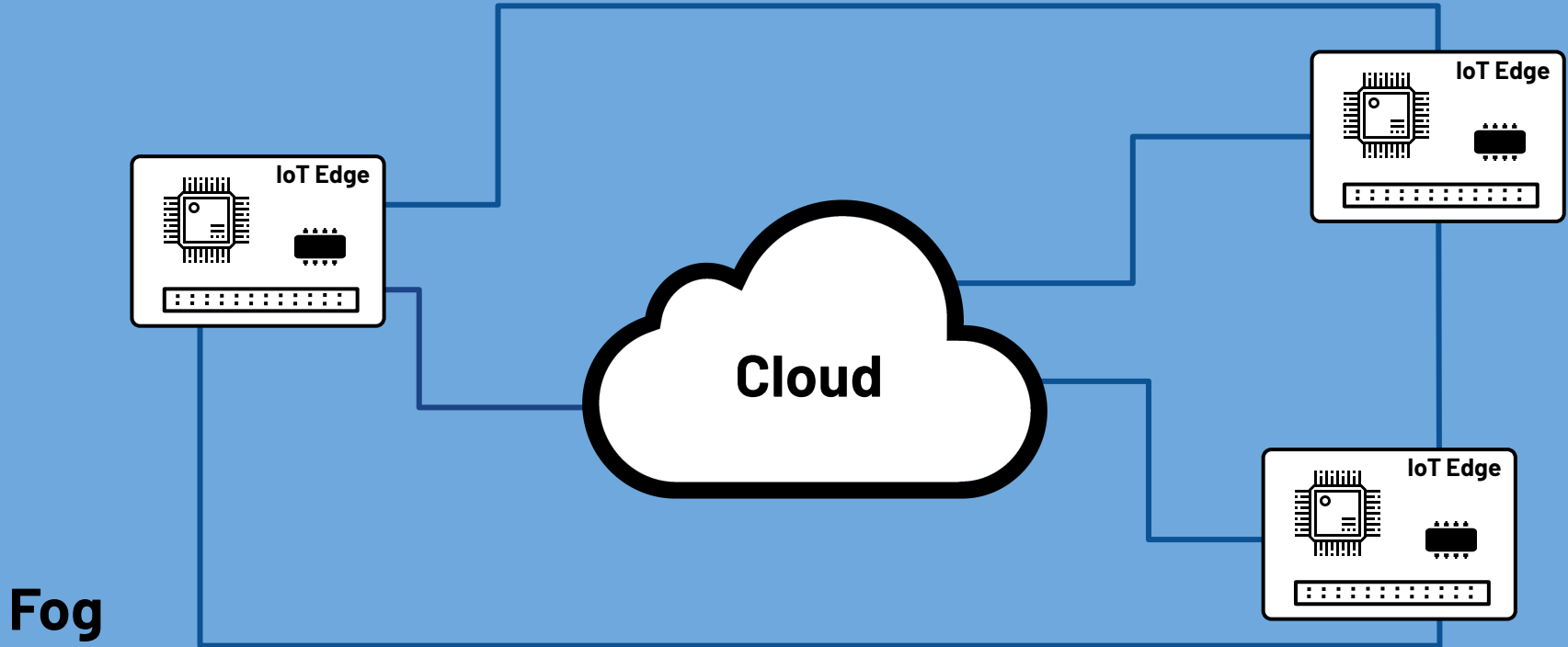
Gateway



Cloud



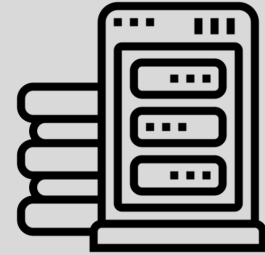
What is Fog Computing?



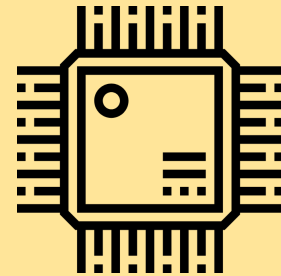
Cloud Computing

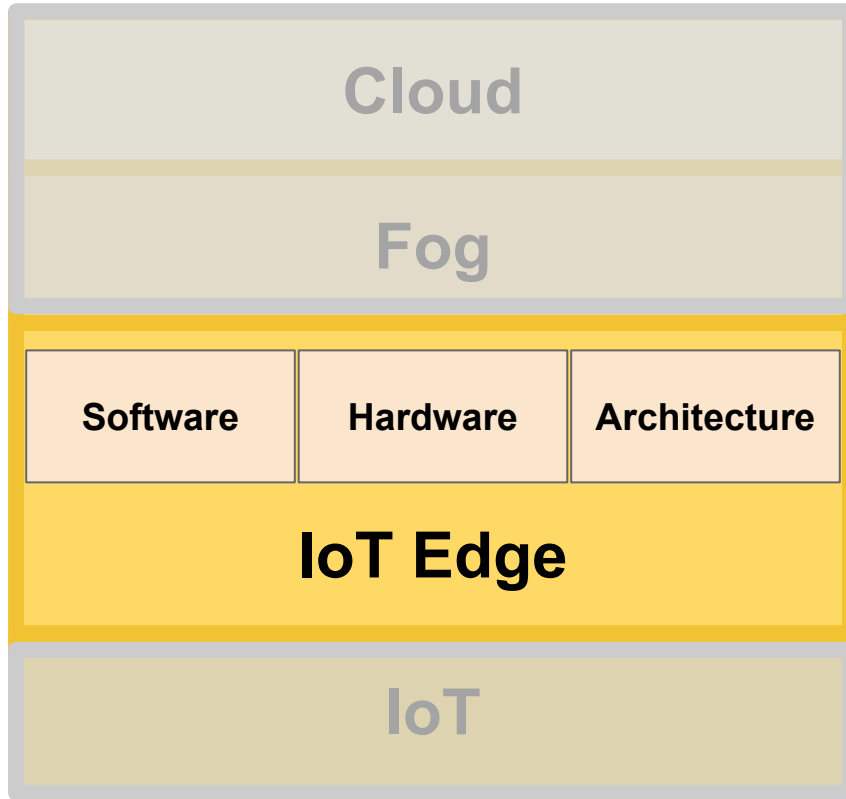


Fog Computing

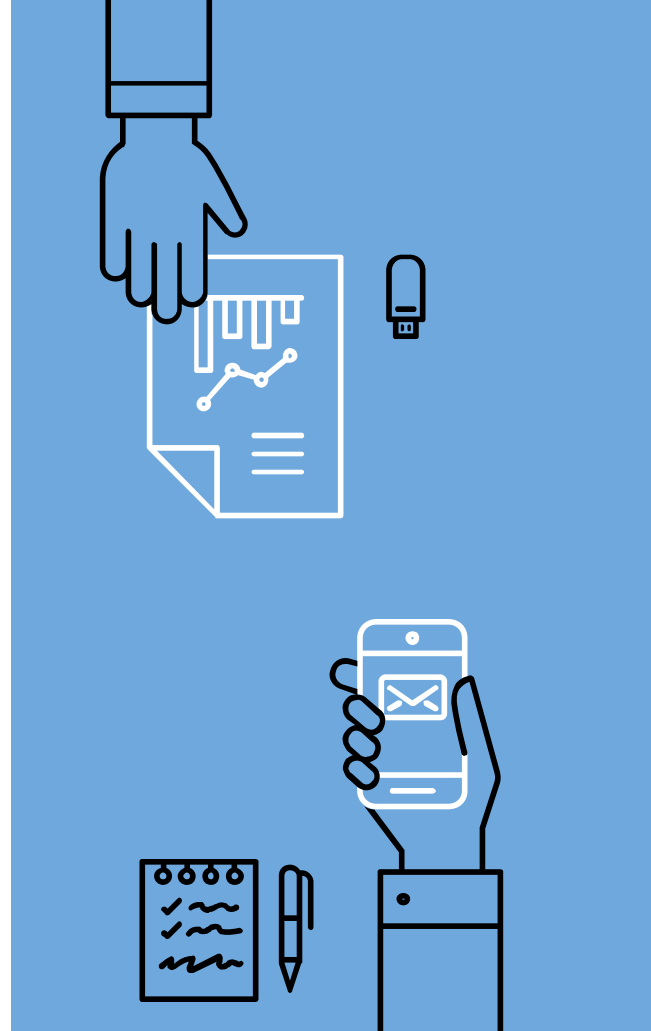
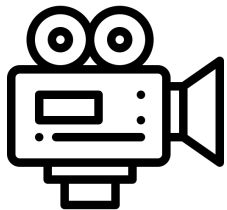


Edge Computing





Real-Life Example



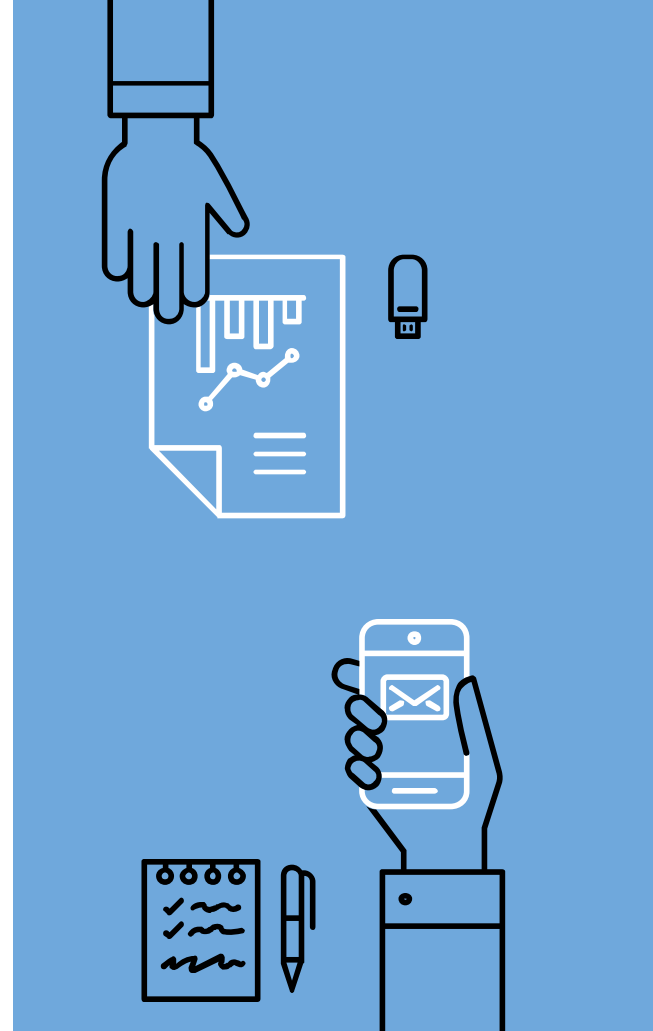
ML-based People Counter

 Privacy Concerns

 Performance

 Latency

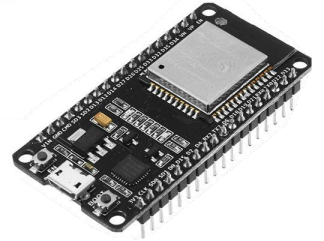
 Cost



Edge Computing Devices

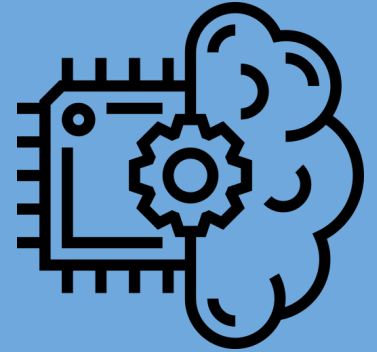
- Azure Sphere
- AWS FreeRTOS
- Google Coral

Raspberry Pi, Banana Pi, Intel Edison ++



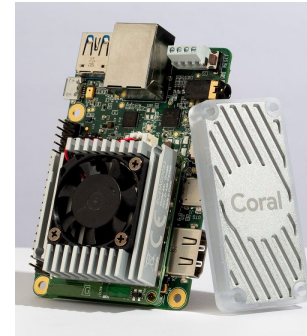
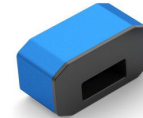
Edge AI Computing

Edge Computing + AI =



AI Accelerators

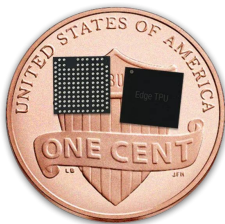
- NVIDIA Jetson
- Intel Movidius Myriad
- Google Edge TPU



Google Coral Dev Board



Not all TPUs are alike...



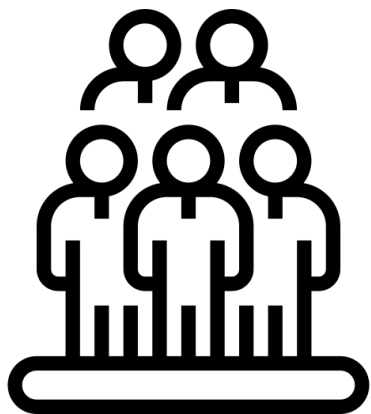
USB-C Port



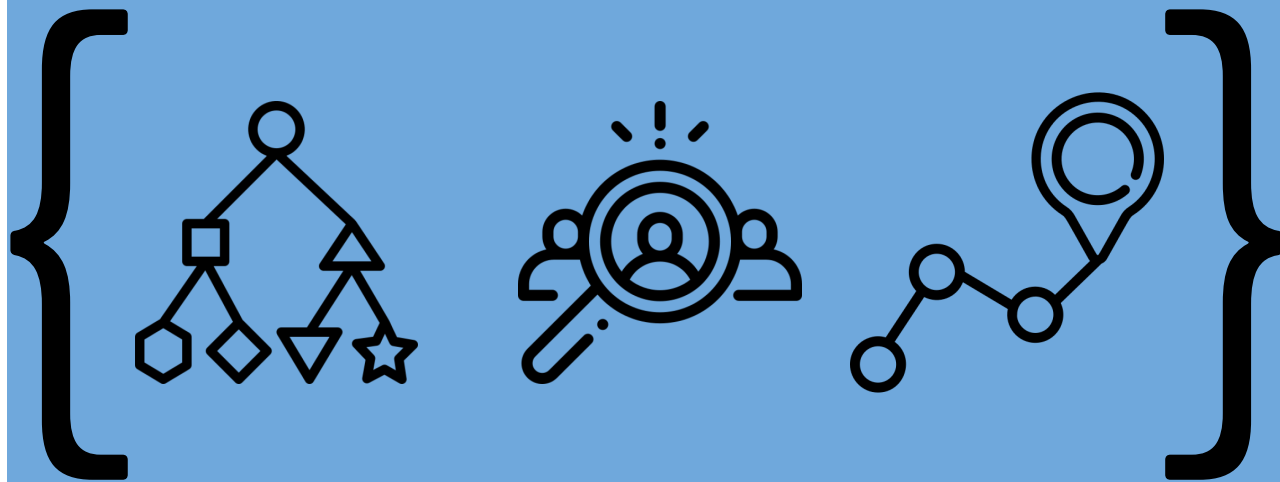
Grown-up Person



<https://cloud.google.com/tpu/>



People
Counter



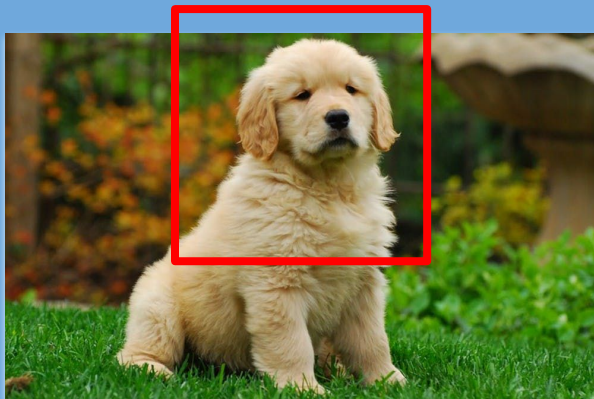
Classification

Detection

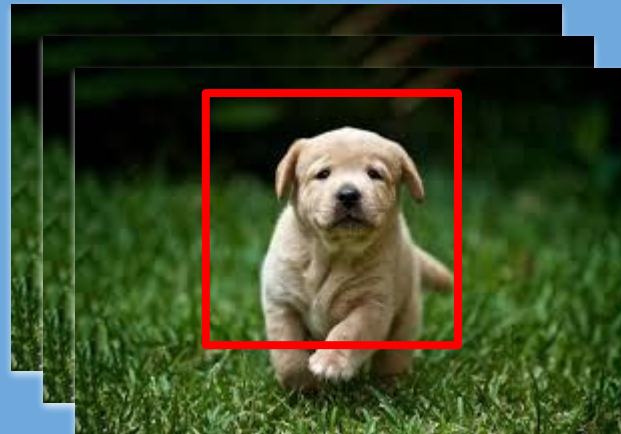
Tracking



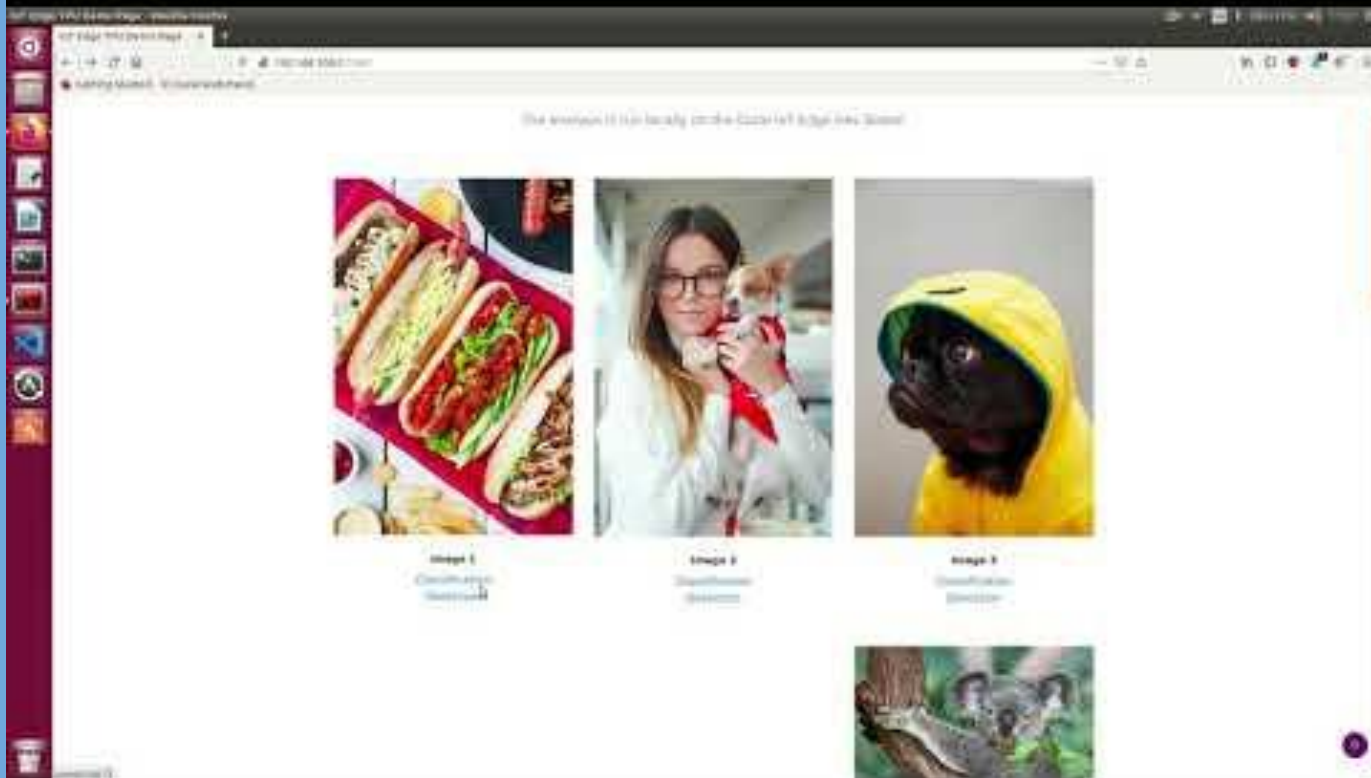
Classification



Detection



Tracking



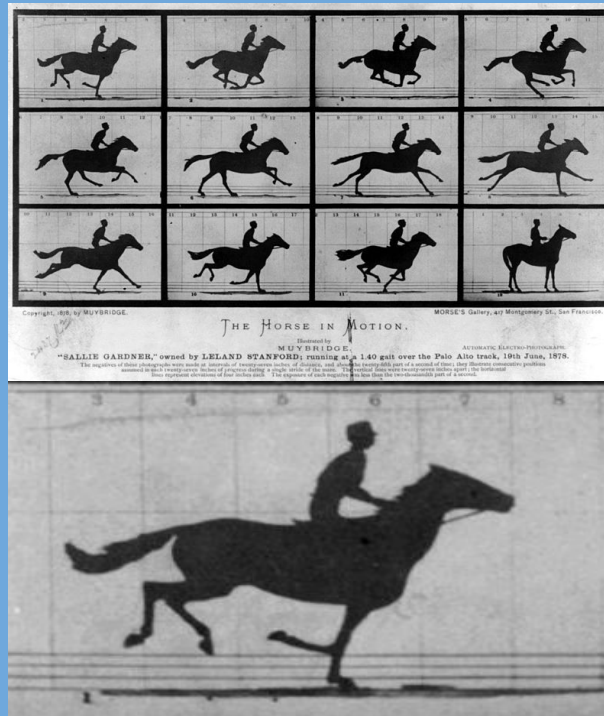
Detection vs. Tracking

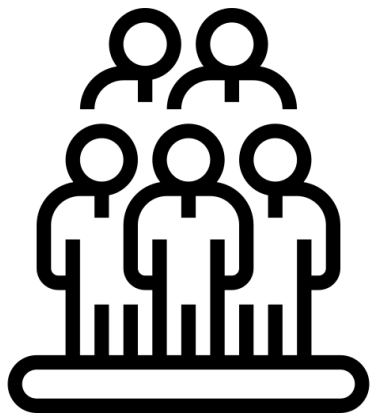
- Video -> stream of images (fps)
- Detect an object in each image.
- Tracking = correlating objects between frames

Lots of images!

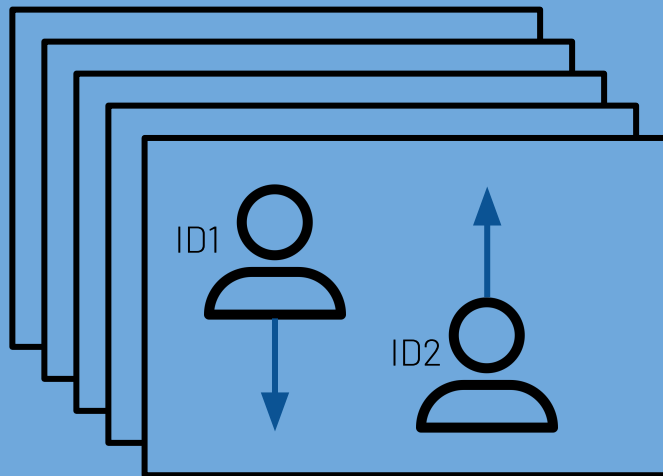
ML is CPU intensive, but not for TPU!



Quite CPU intensive and tricky!

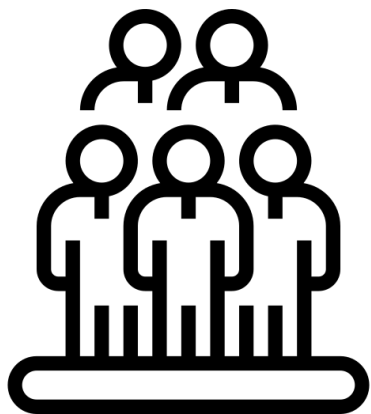




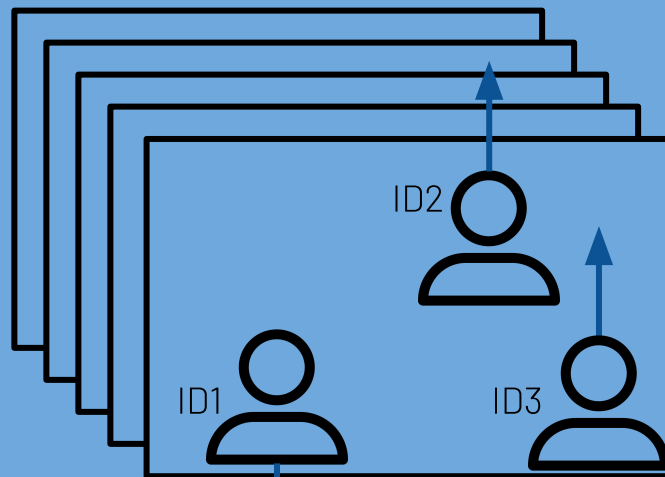
People Counter





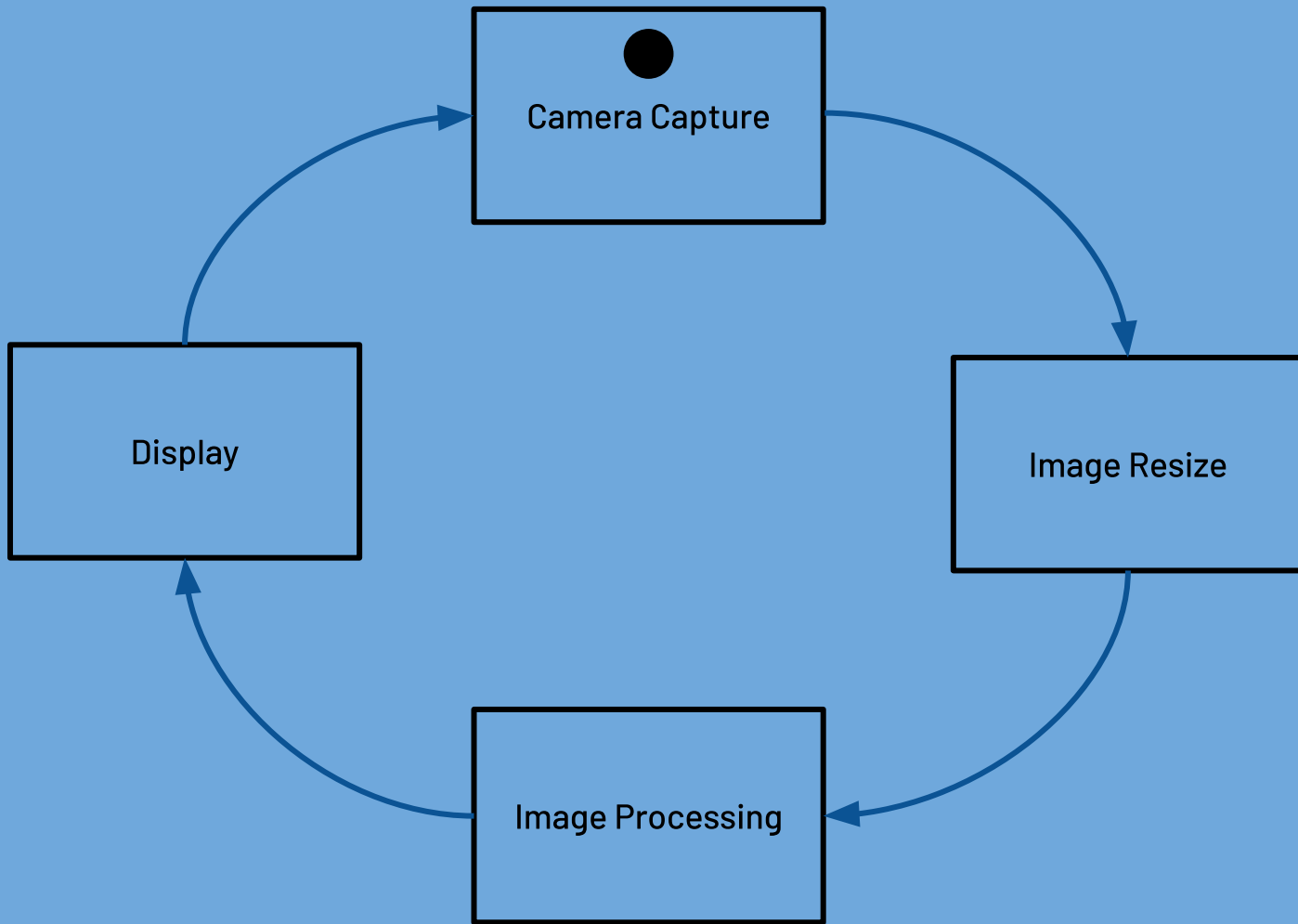
 10 seconds }
 30 fps } = 300 frames



People
Counter



 10 seconds }
 30 fps } = 300 frames







Coral

Edge TPU Performance Demo

The video below demonstrates the real-time processing power of the Edge TPU by running a MobileNet SSD model that can identify and classify multiple objects. The footage of the cars is a recording, but the MobileNet model is executing in real-time on your Coral Dev Board to detect each car indicated with a box (limited to 20 detected cars).

In the terminal where you started the demo, press the W key to switch between running the model on either the Edge TPU or the CPU (quad-core Cortex A73).

```
INFO:edgetpuvision.streaming_server:[192.168.100.52:36448] Stopped.
INFO:edgetpuvision.streaming_server:Number of active clients: 1
INFO:edgetpuvision.streaming_server:[192.168.100.52:36448] Stopping...
INFO:edgetpuvision.streaming_server:[192.168.100.52:36448] Stopped.
INFO:edgetpuvision.streaming_server:Number of active clients: 0
INFO:edgetpuvision.streaming_server:New web connection from 192.168.100.52:36448
INFO:edgetpuvision.streaming_server:Number of active clients: 1
INFO:edgetpuvision.streaming_server:[192.168.100.52:36448] Rx thread finish
ed
INFO:edgetpuvision.streaming_server:[192.168.100.52:36448] Tx thread finish
ed
INFO:edgetpuvision.streaming_server:New web connection from 192.168.100.52:36448
INFO:edgetpuvision.streaming_server:[192.168.100.52:36448] Upgraded to M48 socket
INFO:edgetpuvision.streaming_server:Number of active clients: 2
INFO:edgetpuvision.streaming_server:[192.168.100.52:36448] Stopping...
INFO:edgetpuvision.streaming_server:[192.168.100.52:36448] Stopped.
INFO:edgetpuvision.streaming_server:Number of active clients: 1
INFO:edgetpuvision.streaming_server:[192.168.100.52:36448] stream control 7
INFO:edgetpuvision.streaming_server:[192.168.100.52:36448] breaking client
INFO:edgetpuvision.streaming_server:New web connection from 192.168.100.52:36448
INFO:edgetpuvision.streaming_server:Number of active clients: 1
INFO:edgetpuvision.streaming_server:[192.168.100.52:36448] Rx thread finish
ed
INFO:edgetpuvision.streaming_server:Camera start recording
fileurl: /usr/share/edgetpu_demo/video_stream.asp | qdmax=1 | use_w
=1
I: | gsmc max-size-buffers=1 | h264enc=1 | video/x-h264,atemporal=1,stre
am-format=byte-stream | aacenc max-buffers=1 | aac1 | signal=free | drop=false
| amesh=0 | sync=false
I: | gsmc max-size-buffers=1 | decodebin | yifilterbin | filter=glcolorsync
| video/x-raw,format=RGBA,height=360,width=640 | videoconvert | video/x-raw
,format=RGB,height=180,width=320 | videomx autocrop=true | video/x-raw,bit
rate=360,width=320 | aacenc max-buffers=1 | omit-signal=1 | True | drop=true | name
=signal | sync=false
INFO:edgetpuvision.streaming_server:[192.168.100.52:36448] Tx thread finish
ed
INFO:edgetpuvision.streaming_server:[192.168.100.52:36448] Stopping...
INFO:edgetpuvision.streaming_server:[192.168.100.52:36448] Stopped.
INFO:edgetpuvision.streaming_server:Number of active clients: 1
```

Our Experiences

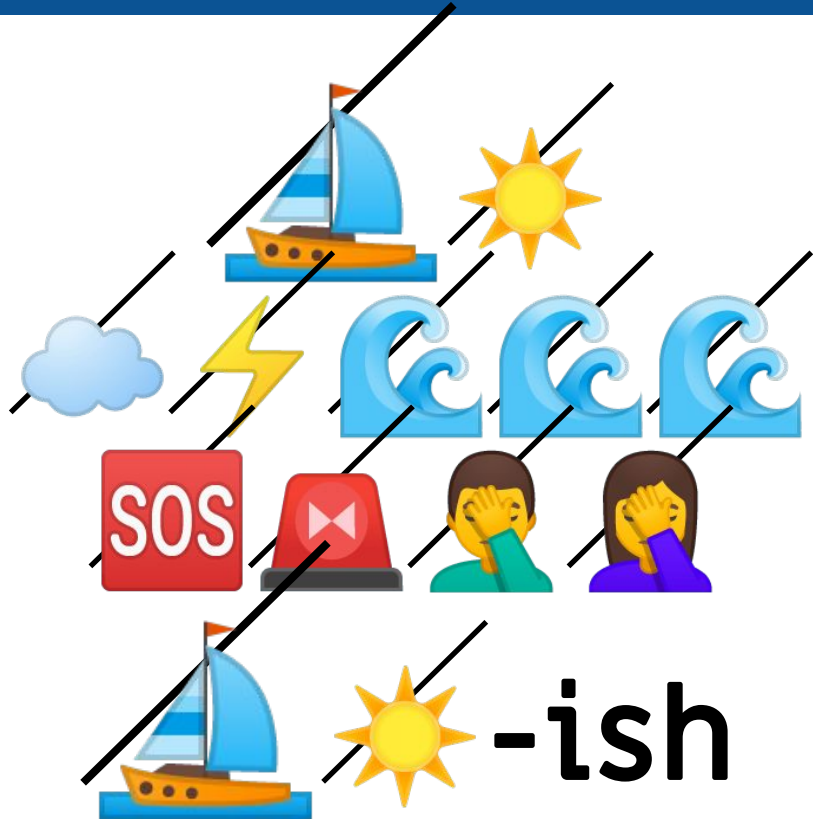
- Various models have difference performance
- Don't put much code in synchronous functions
- Optimize, optimize, optimize

```
top - 10:15:47 up 2:24, 3
Tasks: 161 total, 2 running
%Cpu(s): 24.7 us, 0.9 sy,
KiB Mem : 1007372 total,
KiB Swap: 15707132 total, 15

PID USER PR NI VIRT
6133 root 20 0 4904
801 root 20 0
547 root 20 0
1213 root 20 0
3066 avahi 20
4795 mend 20
```

Når du har 1 GB minne og
lager 16 GB SWAP for at
ting skal kompilere...

Smooth Sailing?



-ish

DEMO

THANK YOU!

Tannaz N. Roshandel

@tannaznvr

Rustam Mehmandarov

@rmehmandarov



UiO • **University of Oslo**



computas