L1 Introduction and Overview

Zonghua Gu 2022

Logistics

- Lecture Times
  - Mon, Wed 15:00-16:30
- Instructors
  - Zonghua Gu, Kalle Prorok, Siyu luan
  - Emails: firstname.lastname@umu.se
- Zoom link:
- https://umu.zoom.us/j/67824756675

My Instruction Style

- No textbook. Will discuss classic techniques as well as latest research advances.
- I provide detailed, self-contained PowerPoint slides.
  - All exam questions will come from these slides. Some slides may be verbose for the sake of completeness.
- In-class questions and discussions are welcome. You can either speak up during lecture, or type in the chat window. For questions after class, please use the Canvas discussion board so everyone can see the discussions.
- Lecture videos are recorded and available in UmU Play, so in-class attendance is not mandatory.

Lecture Schedule (Tentative)

- I will put lecture materials here instead of on Canvas. Please bookmark this link
  - https://guzonghua.github.io/saav/ (under construction)
  - Since I may update slides slightly after each class, it is more convenient to put them here
- Last year's materials available on https://guzonghua.github.io/saav2021/ and YouTube
  - Contents will be updated to have broader scope and less math

Labs and Grades

- Labs
  - Lab1 in W3-4. Adversarial attack on CNN for traffic sign classification
  - Lab2 in W7-8. PID control
  - Lab3 in W9-10. Planning for Highway Driving with DQN RL
- We will keep the projects' computing demands low, so you can use Google Colab, or work on your own computer without powerful GPUs. (Programming language is Python.)
- We will keep the programming workload relatively low, e.g., you may be given a semi-complete program, and asked to tune some hyper parameters, or fill in a few lines of missing code (no large-scale coding)
- Grade distribution:
  - Final exam (open-book): 60%
  - Lab sections: 40%

Final Exam Format

- Multiple choice questions, e.g.
  - Which path planning algorithm(s) are guaranteed to find the optimal solution?
  - A. A* algorithm
  - B. Rapidly-exploring Random Tree (RRT)
  - C. Probabilistic Roadmap (PRM)
  - D. All of them
  - E. None of them
- Simple calculation questions, e.g.,
  - Convolutional Neural Networks I
  - Input volume: 56×56×64 ($W1$=$H1$=$N1$=56,$D1$=64). 32 1×1×64 filters (K=32,F=1) w. stride S=1, no pad P=0. Show the formulas and calculation process.
  - 1) Calculate the dimensions of the output volume, including spatial size and depth.
  - 2) Calculate the total number of parameters, including weights and biases.

Pass or Fail?

- In the past, the vast majority of students pass the course, if they put in reasonable effort
- The course workload is not very high. Most students manage it quite well
  - Use the anonymous feedback link to provide comments on course pace, level of difficulty, etc.

Today’s Agenda

- I will give a broad review of the major issues involved in AD.
  - Background
  - Sensors and perception
  - HD maps
  - Hardware platforms
  - Software platforms
  - V2X
  - Ethical Issues

Autonomous Vehicles (AVs)

- Can refer to any type of Autonomous Mobile Robot.
  - Not just Self-Driving Cars (SDCs)
- Many techniques for SDCs covered in this course are generally applicable to other types of AVs.

Self-Driving Cars

Drones

Warehouse Robots

Indoor-Cleaning Robots

Why AD?

- Reduced traffic accidents and fatalities
  - In the USA: in 2019, an estimated 38,800 people lost their lives to car crashes. About 4.4 million people were injured seriously enough to require medical attention in crashes.
- Reduced congestion and pollution
- More productive time spent on the road
- Autonomous Mobility-on-Demand (AMoD) with a fleet of AVs
  - Low-cost, safe and efficient mode of transportation that may make vehicle ownership obsolete.
  - The dream of Uber (and many other companies)

Advanced Driver Assistance (ADAS)

- Since the initial introduction of Cruise Control in 1948, ADAS functions are increasingly prevalent in modern vehicles.
  - Adaptive cruise control (ACC), Anti-lock braking system, Collision avoidance system (Pre-crash system), Driver Monitoring System (DMS), Electronic Stability Control (ESC), Forward Collision Warning (FCW), Lane Departure Warning (LDW), Lane Change Assistance, Surround View…

https://en.wikipedia.org/wiki/Advanced_driver-assistance_systems

DARPA Grand Challenge (2004)

- Held in the Mojave Desert region of the USA, along a 150-mile route.
- None of the robot vehicles finished the route. Carnegie Mellon University's Red Team and car Sandstorm (a converted Humvee) traveled the farthest distance, completing 7.32 mi of the course before getting hung up on a rock after making a switchback turn.
- No winner was declared, and the cash prize was not given.

DARPA Grand Challenge (2005)

- Vehicles passed through three narrow tunnels and negotiated more than 100 sharp left and right turns.
- Five vehicles successfully completed the 132 mi course. Stanford's Stanley won the $2M top prize.

Vehicle	Team Name	Team Home	Time Taken (h:m)	Result
Stanley	Stanford Racing Team	Stanford University, Palo Alto, California	6:54	First place
Sandstorm	Red Team	Carnegie Mellon University, Pittsburgh, Pennsylvania	7:05	Second place
H1ghlander	Red Team		7:14	Third place
Kat-5	Team Gray	The Gray Insurance Company, Metairie, Louisiana	7:30	Fourth place
TerraMax	Team TerraMax	Oshkosh Truck Corporation, Oshkosh, Wisconsin	12:51	Over 10-hour limit, fifth place

DARPA Urban Challenge (2007)

- Held at the site of the now-closed George Air Force Base. The course involved a 60 mi urban area course, to be completed in less than 6 hours. Rules included obeying all traffic regulations while negotiating with other traffic and obstacles and merging into traffic.
- CMU's Boss won the $2M top prize, and Stanford's Junior won the $1M second prize.
- The 3 Grand Challenge races jump-started the Self-Driving Car industry. Faculty and students from winning teams such as Stanford and CMU later became leaders in SDC projects at companies like Google/Waymo and Uber and numerous startups.

Team Name	ID#	Vehicle	Type	Team Home	Time Taken (h:m:s)
Tartan Racing	19	Boss	2007 Chevy Tahoe	Carnegie Mellon University, Pittsburgh, Pennsylvania	4:10:20
Stanford Racing	03	Junior	2006 Volkswagen Passat Wagon	Stanford University, Palo Alto, California	4:29:28
VictorTango	32[11]	Odin	2005 Ford Hybrid Escape	Virginia Tech, Blacksburg, Virginia	4:36:38
MIT	79	Talos	Land Rover LR3	MIT, Cambridge, Massachusetts	Approx. 6 hours
The Ben Franklin Racing Team	74	Little Ben	2006 Toyota Prius	University of Pennsylvania, Lehigh University, Philadelphia, Pennsylvania	No official time.
Cornell	26	Skynet	2007 Chevy Tahoe	Cornell University, Ithaca, New York	No official time.

Highway vs. City Driving

- Highway driving is perceived as an easier problem than city driving.
  - Has potential of massive displacement of truck driver jobs
  - But traffic merging is tricky and may require human operator assistance

	Highway Driving	**City Driving**
Travel Speed	High	Low to medium
Traffic Volume	High	Medium to high
Number of Lanes	Large (6-8)	Small (2-4)
Others	Entry and exit points for traffic merging	Many intersections with traffic lights

Brief History of SDCs

- 2004/05 — DARPA Grand Challenges
- 2007 — DARPA Urbane Challenge
- 2009 — Google begins AV project
- 2012 — Google's AV passes a 14-mile driving test in Nevada
- 2013 — Mercedes and Infiniti produce cars with radar sensors and some AD features
- 2013 — NHTSA releases initial policy on AVs
- 2015 — Tesla releases Auto-Pilot
- 2015 — Uber hires 40 CMU robotics researchers to work on AVs; Ford begins testing its AVs.
- 2016 — Major acquisitions and partnerships (GM and Cruise Automation; GM and Lyft; Toyota and Jaybridge Robotics; Uber and Volvo)
- 2016 — NHTSA issues guidelines for testing and deployment of AVs
- 2017 — NHTSA issues revised safety guidelines for AVs

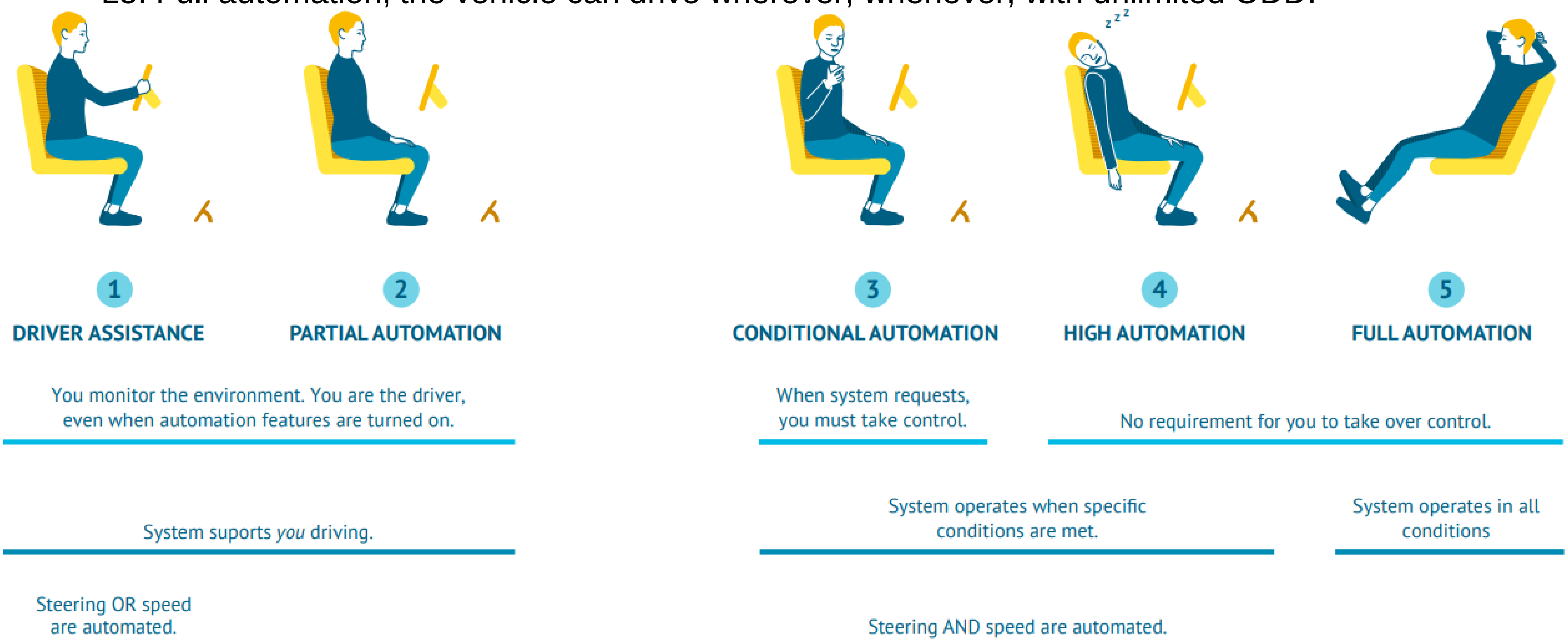
https://www.mcca.com/wp-content/uploads/2018/04/Autonomous-Vehicles.pdf

Operational Design Domain (ODD)

- The ODD defines the conditions under which a vehicle is designed to function and is expected to perform safely. The ODD includes (but isn’t limited to) environmental, geographical, and time-of-day restrictions, as well as traffic or roadway characteristics.
  – e.g., an autonomous freight truck might be designed to transport cargo from a seaport to a distribution center 30 Km away, via a specific route, in day-time only. This vehicles ODD is limited to the prescribed route and time-of-day, and it should not operate outside of it

Five Levels of Automation

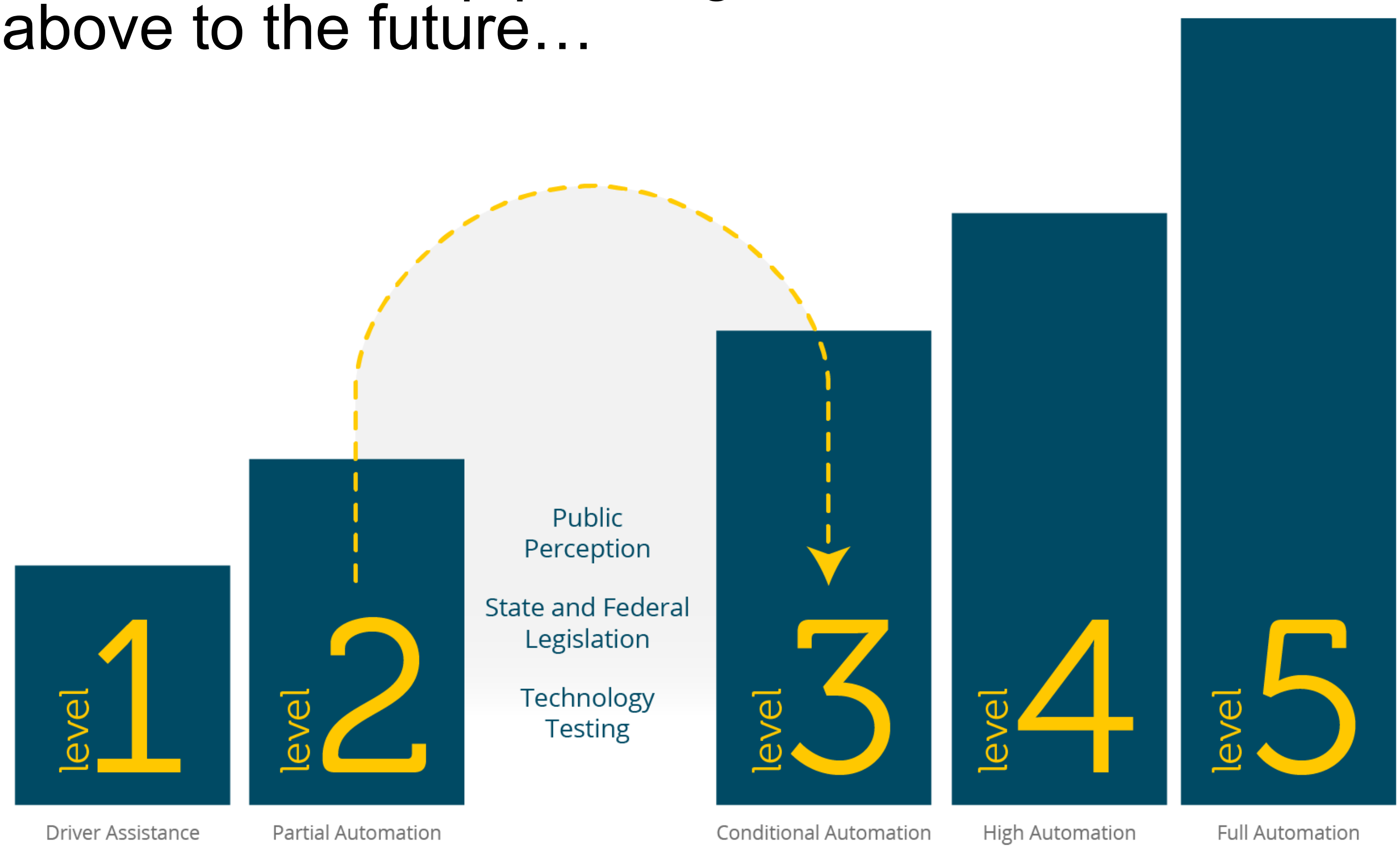
- L1: ADAS features that either control steering or speed to support the driver.
- L2: both steering and acceleration are simultaneously handled by AD system. The human driver still monitors the environment and supervises the support functions.
- L3: Conditional automation: the system can drive without the need for a human to monitor and respond. However, the system might ask a human to intervene, so the driver must be able to take control at all times.
- L4: These systems have high automation and can fully drive themselves under certain conditions. The vehicle won’t drive if not all conditions are met.
- L5: Full automation, the vehicle can drive wherever, whenever, with unlimited ODD.



https://www.wevolver.com/article/2020.autonomous.vehicle.technology.report

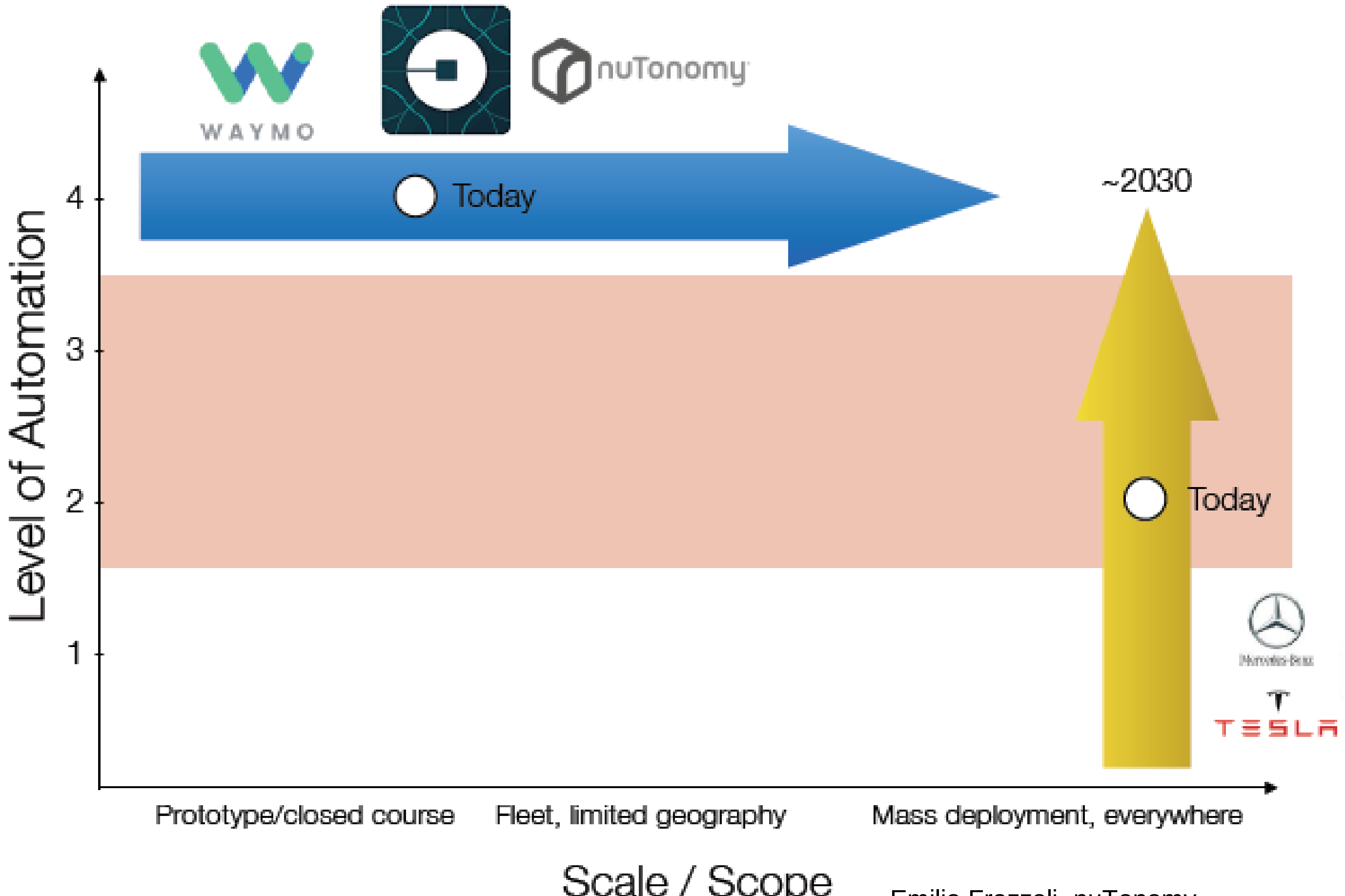
State of the Art

- Current commercial products are at most L2 (e.g., Tesla Autopilot)
- L2 to L3 is perceived to be a giant leap
- Automakers keep pushing the timeline of L3 and above to the future…



Two Different Paths to L4/5

- Tesla starts from L2 and mass deployment, and gradually moves to L4/5.
- Waymo, nuTonomy…starts from L4 in limited ODD, and gradually expands deployment



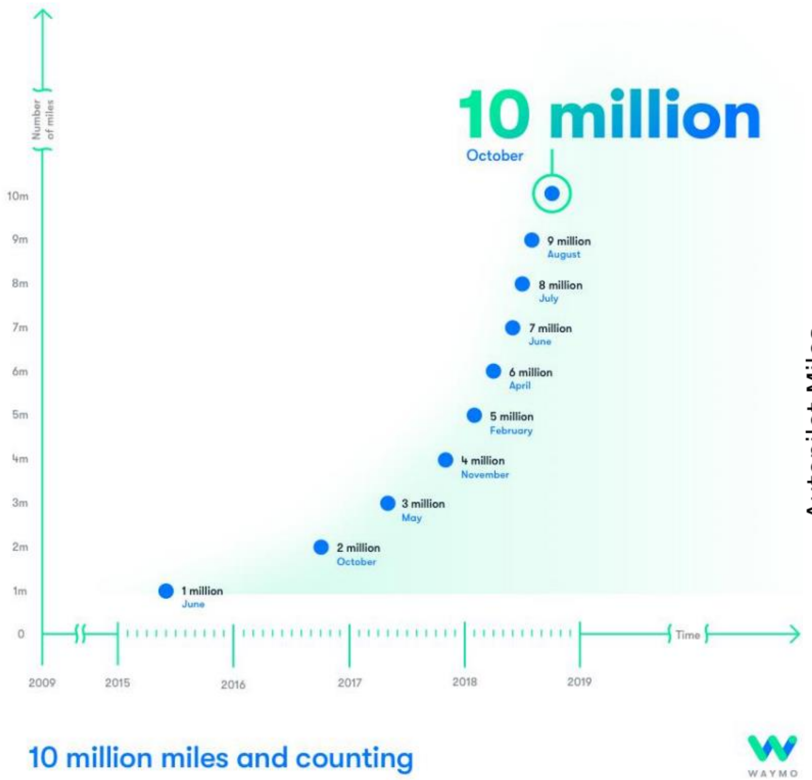
AD Safety Evaluation Metric: Miles Driven?

- Not All Miles are Equal.
  - Driving conditions may be dramatically different.
  - Companies may be incentivized to avoid difficult driving conditions.

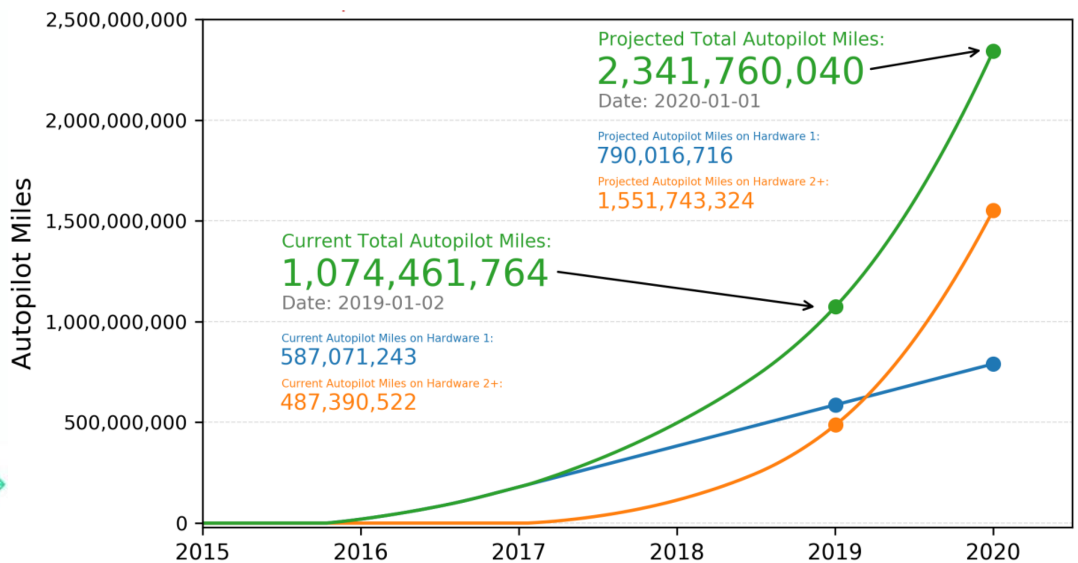
Miles driven here

Not the same as here

Total Miles Driven

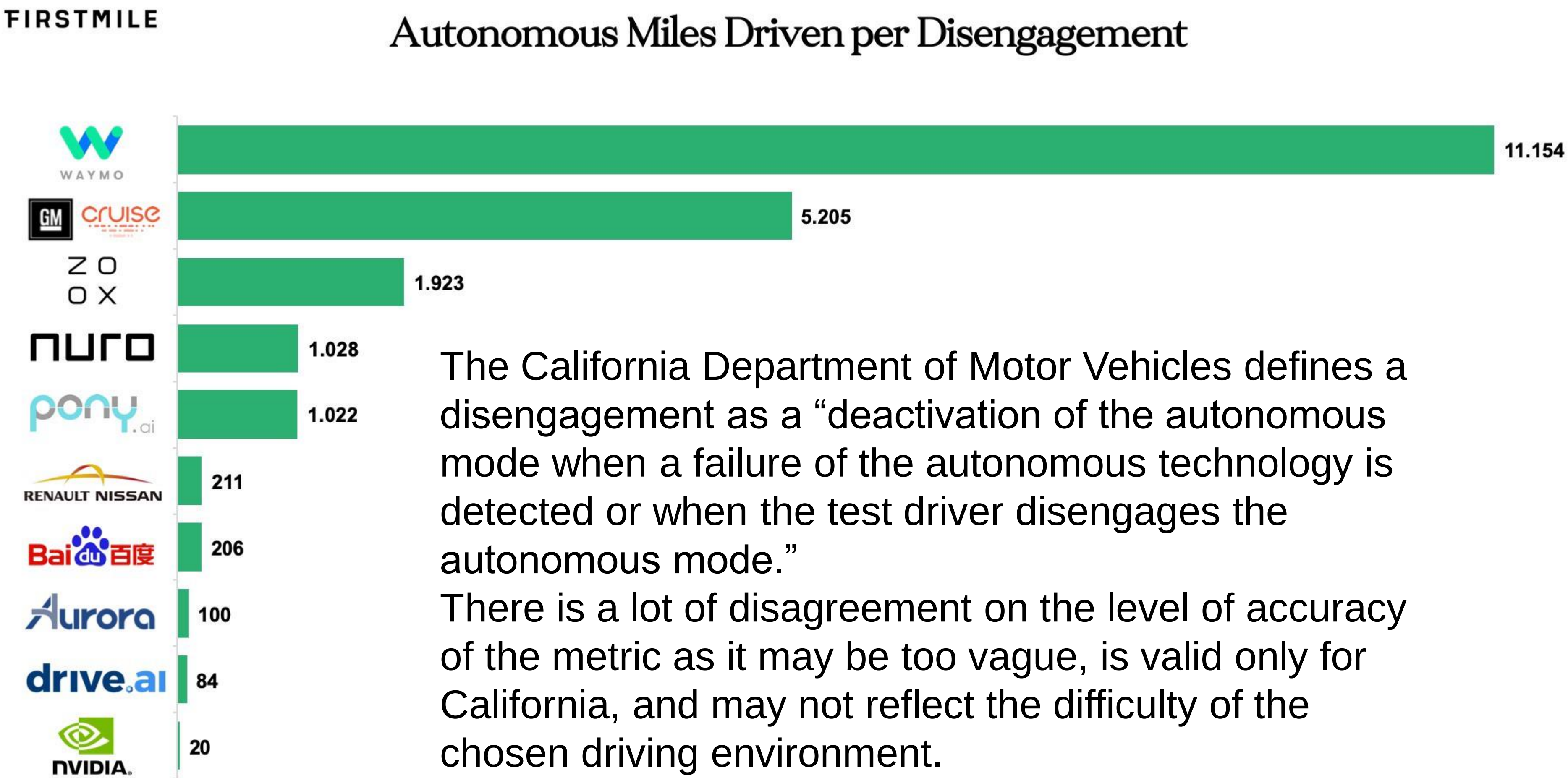


Waymo Reaches 10 Million Miles



Tesla Autopilot Reaches 1 Billion Miles in 2019

Miles Driven per Disengagement (2018)



The California Department of Motor Vehicles defines a disengagement as a “deactivation of the autonomous mode when a failure of the autonomous technology is detected or when the test driver disengages the autonomous mode.”

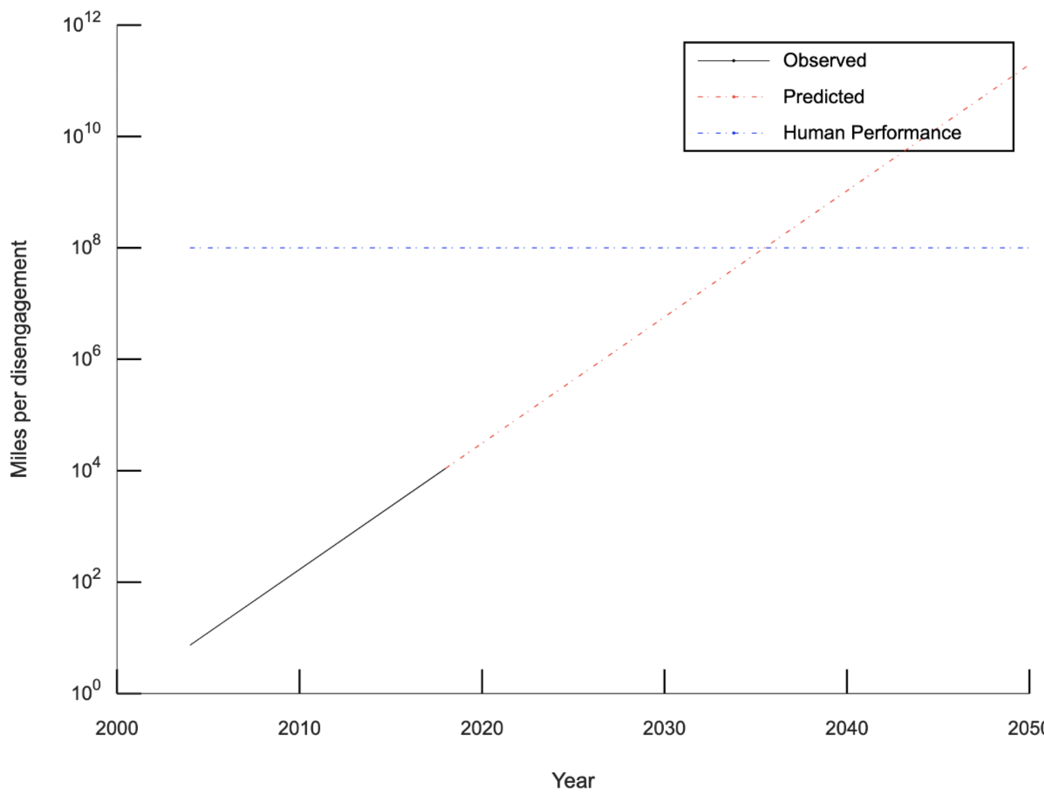
There is a lot of disagreement on the level of accuracy of the metric as it may be too vague, is valid only for California, and may not reflect the difficulty of the chosen driving environment.

Sources: California Department of Motor Vehicles Disengagement Reports February 2019

Decoding the Autonomous Driving Landscape July 2019 | Firstmile | www.firstmile.de

“The Moore’s Law for Self-Driving Vehicles”, Edwin Olson, CEO of May Mobility

- Moore’s law in the semiconductor industry says that “the number of transistors on a chip doubles approximately every 18 months”, i.e., w. exponential growth rate
- Can we have a Moore’s law for AD? “The number of miles between disengagements will double approximately every 16 months.”
  - Between human performance (10^8 miles per fatality) and the best-reported self-driving car performance (10^4 miles per disengagement) is a gap of 10,000x. Even with performance doubling every 16 months, it will reach human levels of performance in 2035.



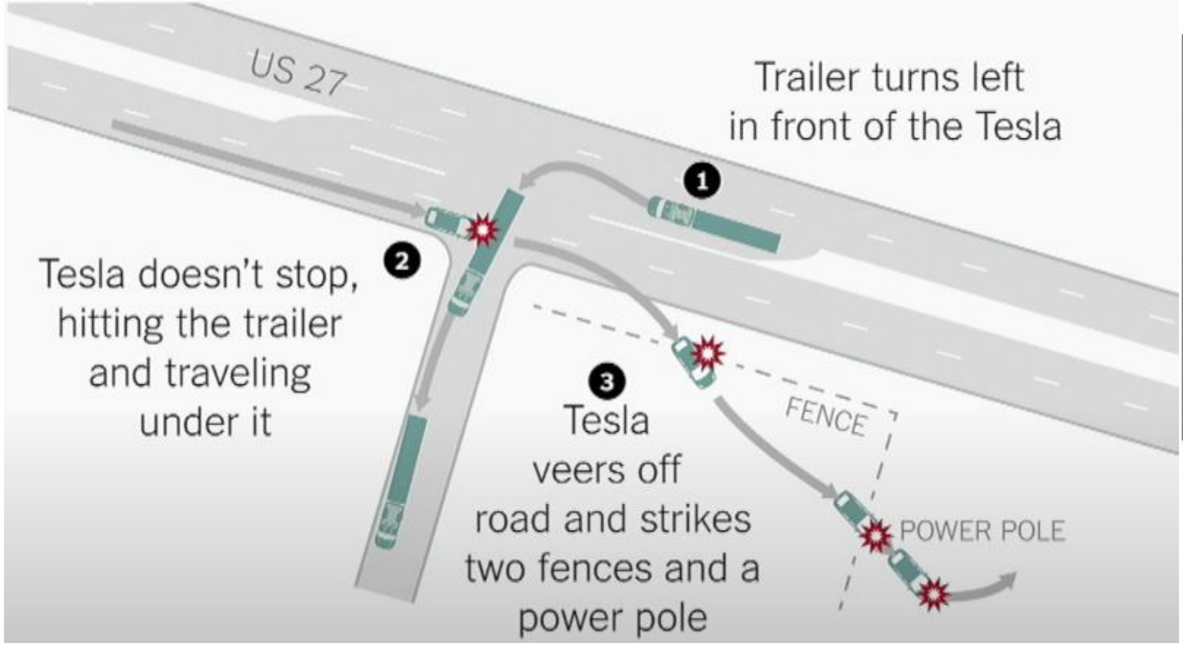
https://medium.com/may-mobility/the-moores-law-for-self-driving-vehicles-b78b8861e184

When will AD Really Arrive?

- Chris Urmson, co-founder and CEO of Aurora:
  - “In 5 years” - 2009
  - “In 5 years” - 2012
  - “In 5 years” - 2015
  - “In 5 years” - 2018

The Tesla Fatality in May 2016

- The Tesla Model S (L2) was driving 74 mph on the highway when it was struck by a semitruck
- The driver's hands were off the steering wheel for a total of 37 minutes during the 37.5 minutes of time the car was in Autopilot, despite repeated visual warnings
- Tesla: "Neither Autopilot nor the driver noticed the white side of the tractor trailer against a brightly lit sky, so the brake was not applied."
  - A failure of computer vision algorithm for object detection; maybe a lidar could have prevented the accident.
- Tesla used to use Mobileye's hardware platform EyeQ, but they broke up after the accident, and Tesla started to develop its hardware platform FSD.



The Tesla Fatality in Mar 2018

- In 2018, a man died from a high-speed crash because his Tesla Autopilot system steered the car into a median on Highway 101 in Mountain View, CA.
- NTSB's investigation report, released in Feb 2020, lists 23 findings that enumerate all the factors that contributed to the fatal collision.
  - Limitations on Tesla's Autopilot Lane-Keeping Assistance (LKA) caused the vehicle to veer into the median and failed to provide an alert to the driver in the seconds leading to the crash.
  - The collision avoidance system was not designed to detect a crash attenuator, which resulted in a severe crash in which the automatic braking and collision warning systems failed to activate.
  - A failure of computer vision algorithm for lane tracking.
- Tesla Autopilot 2 almost crashes Into Barrier (similar to this crash)
  - https://www.youtube.com/watch?v=TIUU1xNqI8w

A Tesla Model X is surrounded by firefighting foam after crashing and catching fire on Highway 101 in Mountain View in March 2018. An NTSB investigation blamed the car's Autopilot system for steering into the median divide, and said the driver likely failed to react because he was playing a video game. Courtesy of Mountain View Fire Department

https://www.mv-voice.com/news/2020/02/25/ntsb-teslas-autopilot-steered-model-x-into-highway-median-causing-fatal-mountain-view-crash

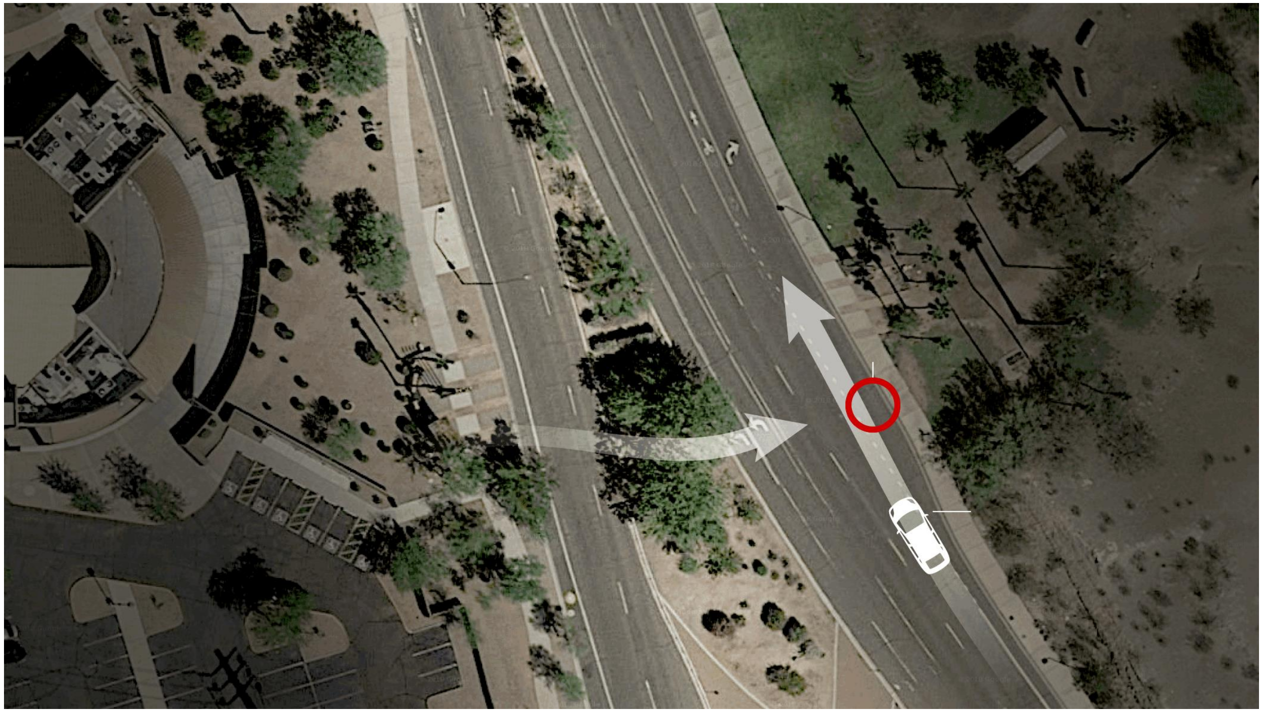
Tesla still cannot recognize white trucks in 2020

- Tesla on autopilot crashes into overturned truck on busy highway in Taiwan
  - https://www.youtube.com/watch?v=X3hrKnv0dPQ

Tesla on autopilot crashes into overturned truck

The Uber Pedestrian Fatality in Mar 2018

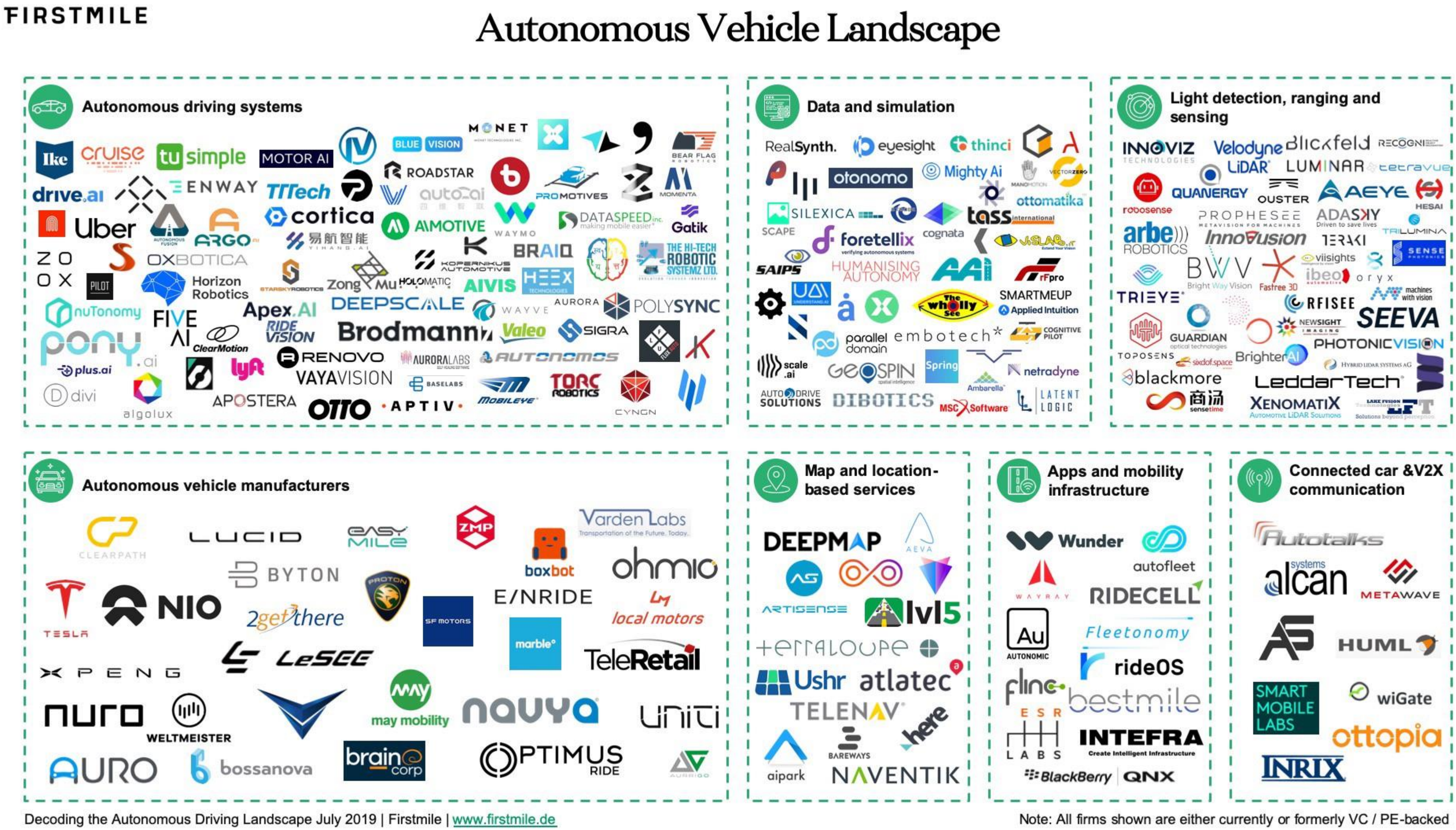
- Police release video of Uber collision that killed pedestrian
  - https://www.youtube.com/watch?v=q7d90ZFhg28
- "The recorded telemetry showed the system had detected Herzberg six seconds before the crash, and classified her first as an unknown object, then as a vehicle, and finally as a bicycle, each of which had a different predicted path according to the autonomy logic."



The Uber Pedestrian Fatality in Mar 2018

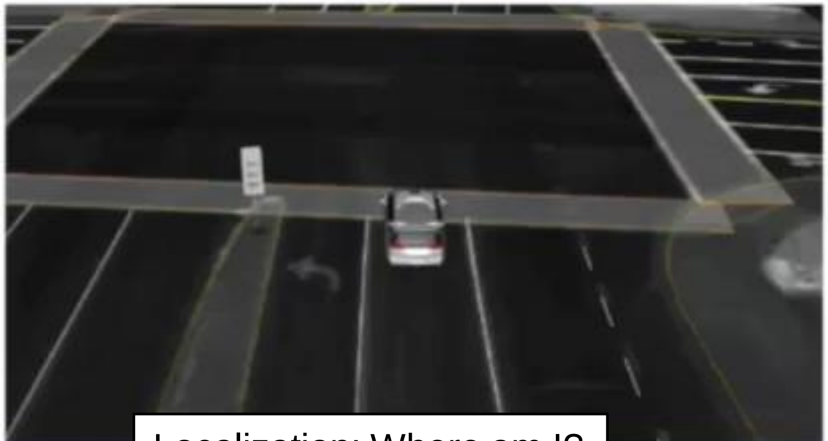
- The AV (L3/4) was equipped with both Lidar and Radar. After the woman was detected on the road (6 sec before)
  - first classified as unknown object
  - then misclassified as a vehicle
  - then a bicycle
- 1.3 sec before, the Volvo system tried to do emergency braking maneuver
  - but Uber had disabled it for testing
- The safety driver was not watching the road moments before the vehicle struck her.
  - It was probably too dark for the driver to see her in time.

AD Landscape Today

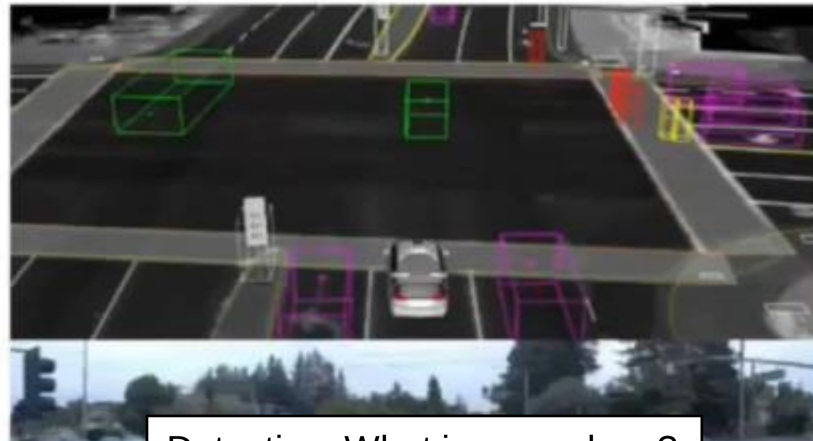


https://medium.com/@firstmilevc/avlandscape-8a21491f1f54

Four Major Tasks of an AV

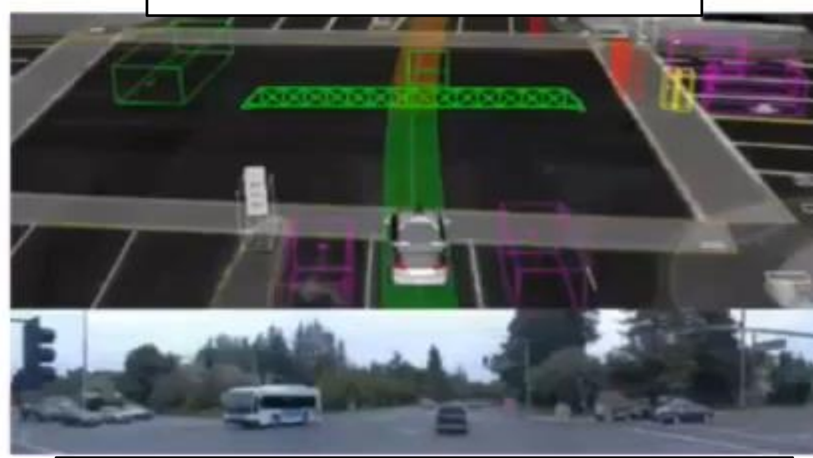


Localization: Where am I?



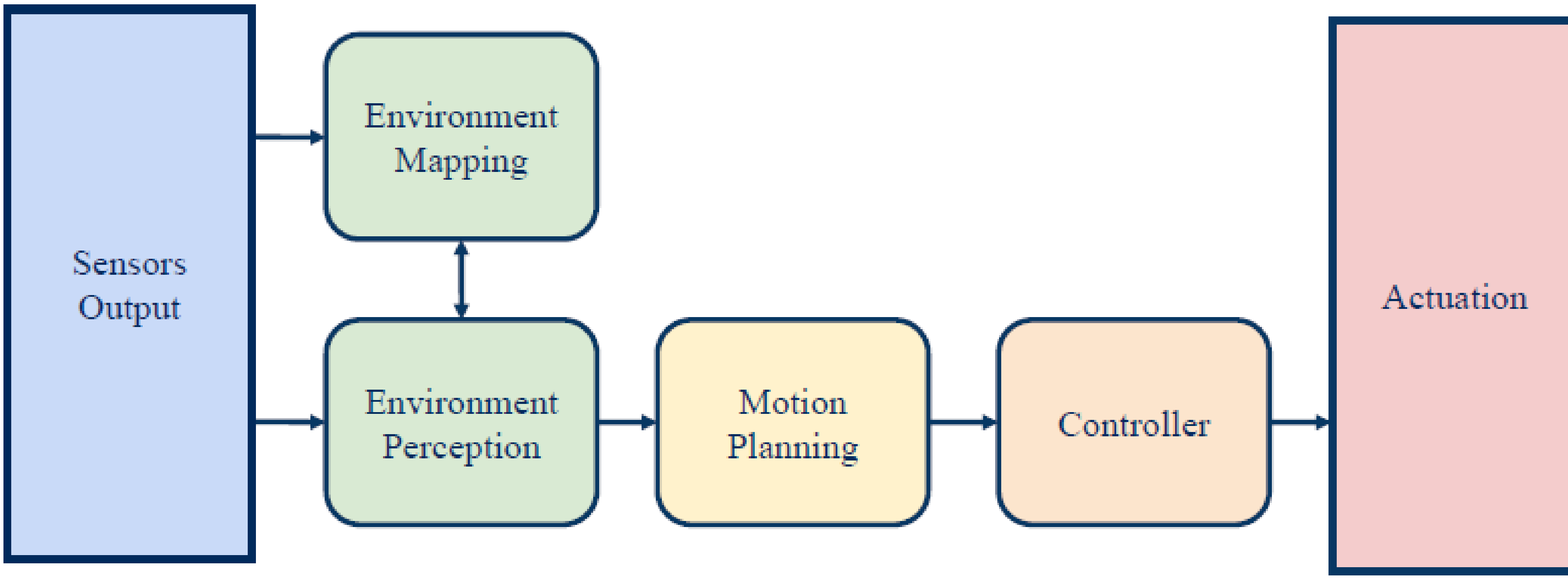
Detection: What is around me?

Prediction: Where are they going?



Planning and Control: Where should I go?

AD Processing Pipeline

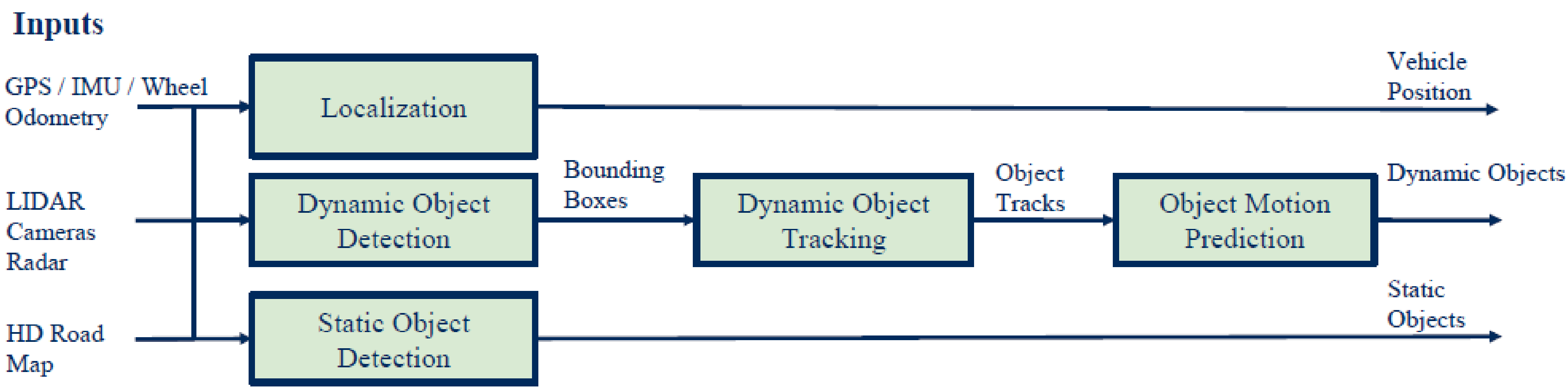


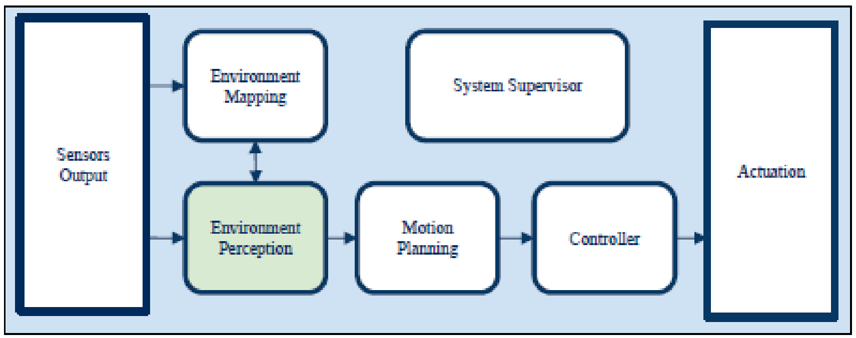
Sensing:
Understanding the surrounding environment, incl. localization, detection, prediction

Planning:
Decision making in the context of other road users

Acting:
Moving the vehicle to follow the planned trajectory

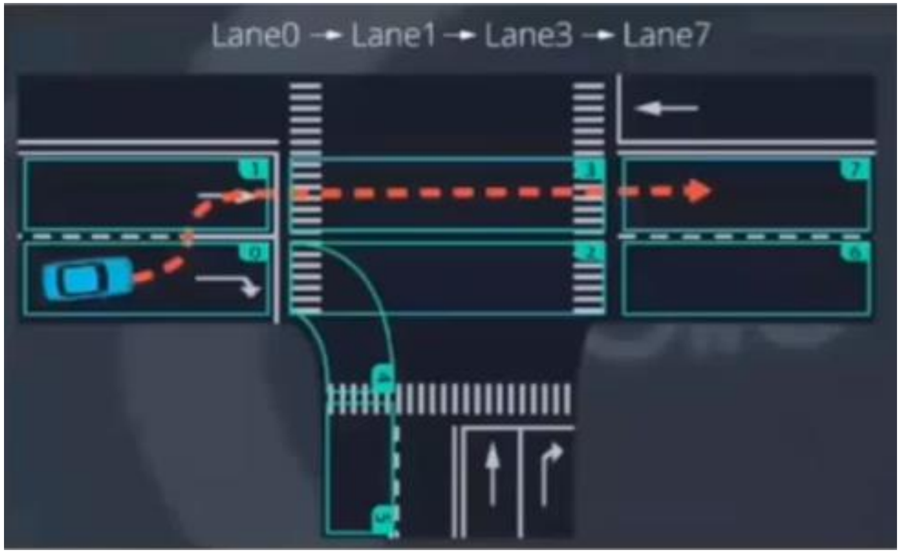
Environment Perception



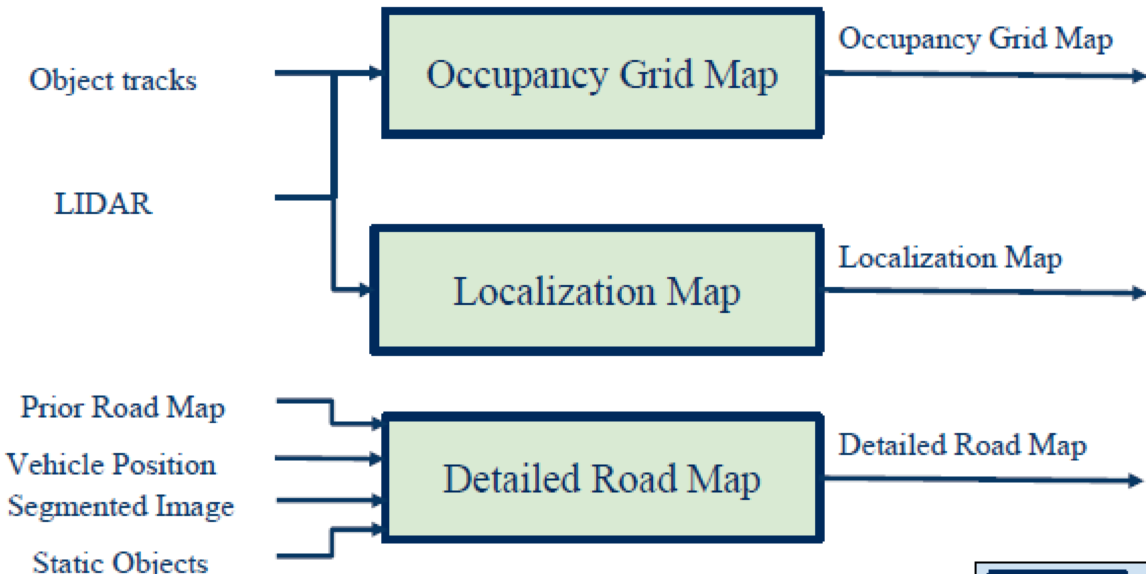


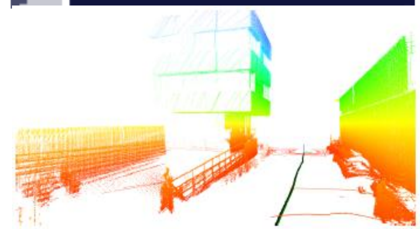
Prediction

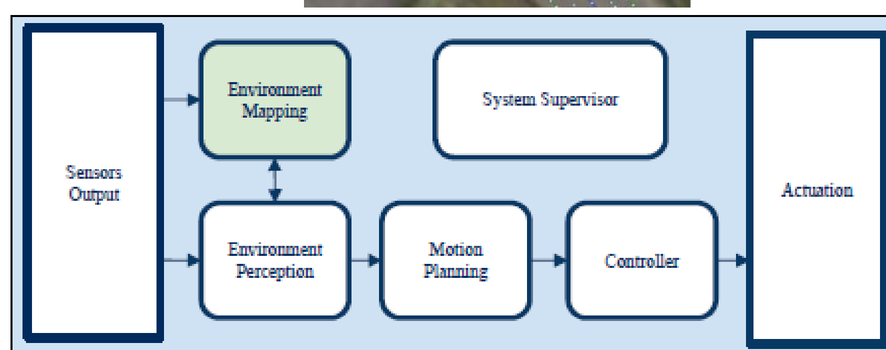
- Based on state
  - Kalman filter
  - Particle filter
- Data-driven
  - ML-based classification
- Pedestrian intention prediction
  - Based on visual cues such as pose, etc.
  - Very difficult problem



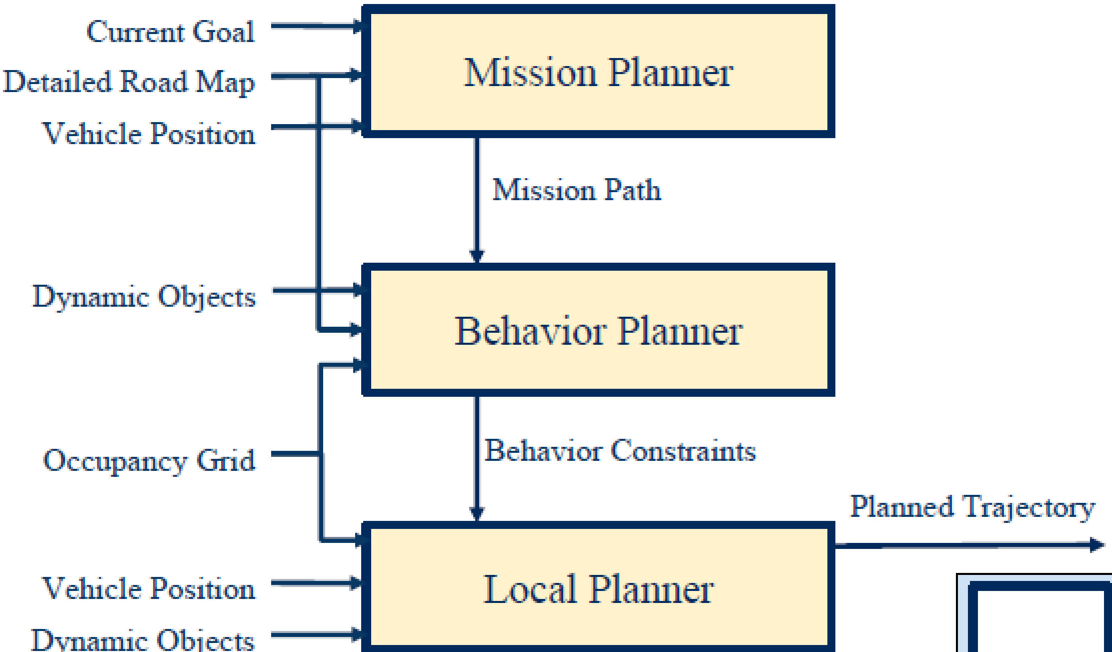
Environmental Maps

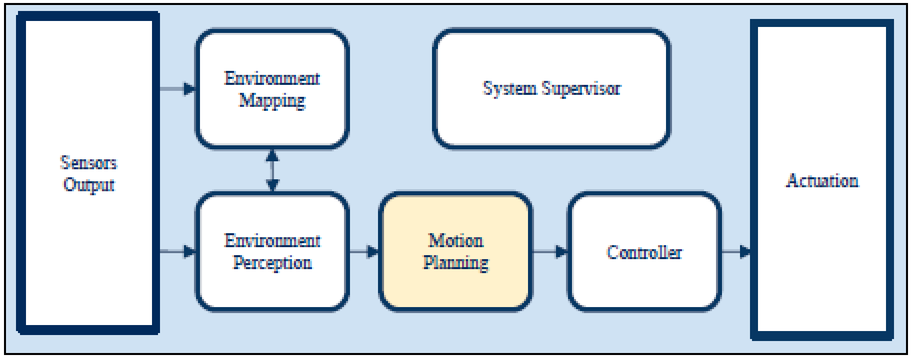






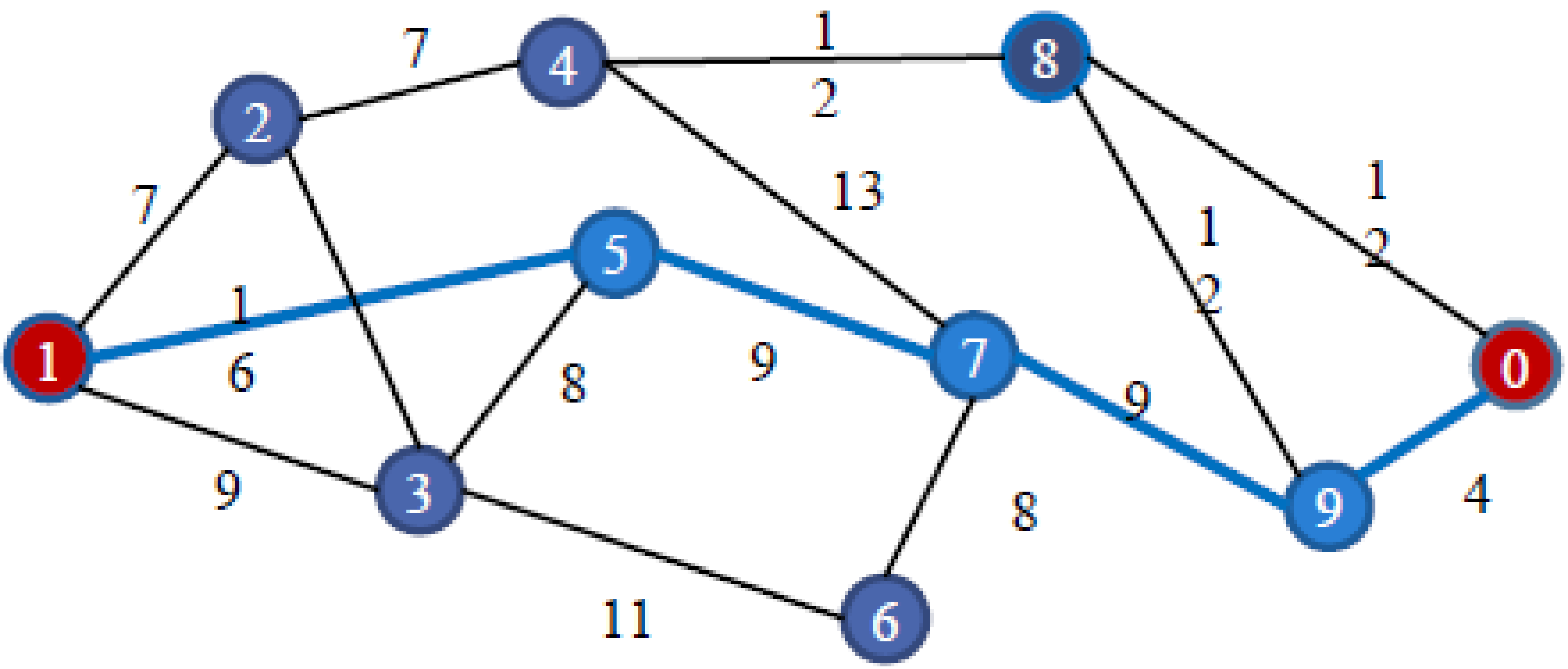
Motion Planning





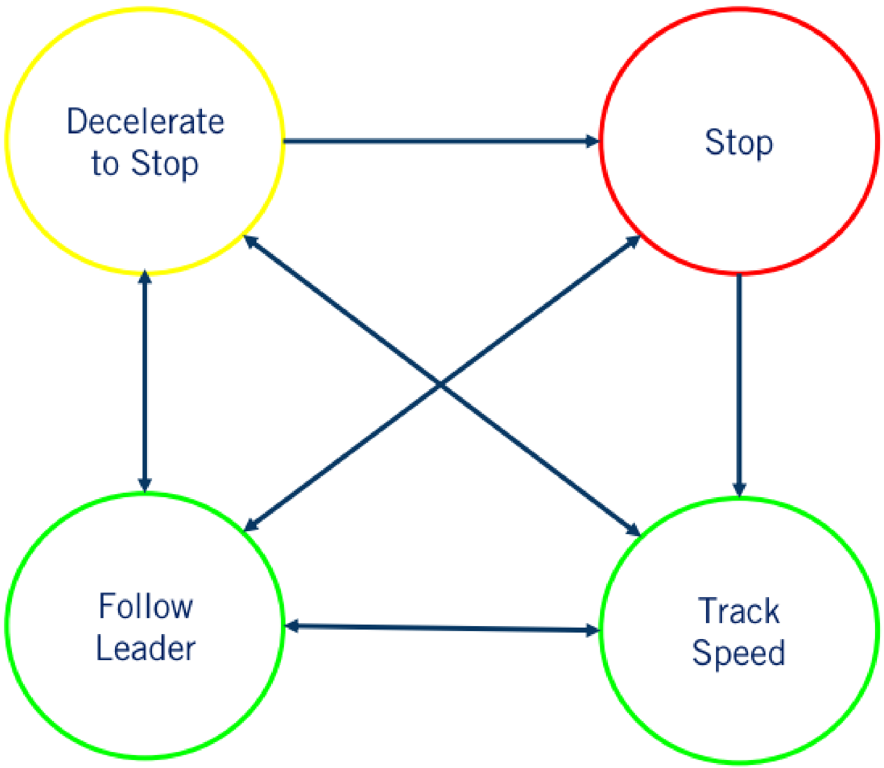
Mission Planner

- Use graph search to find a path from source to destination on the map



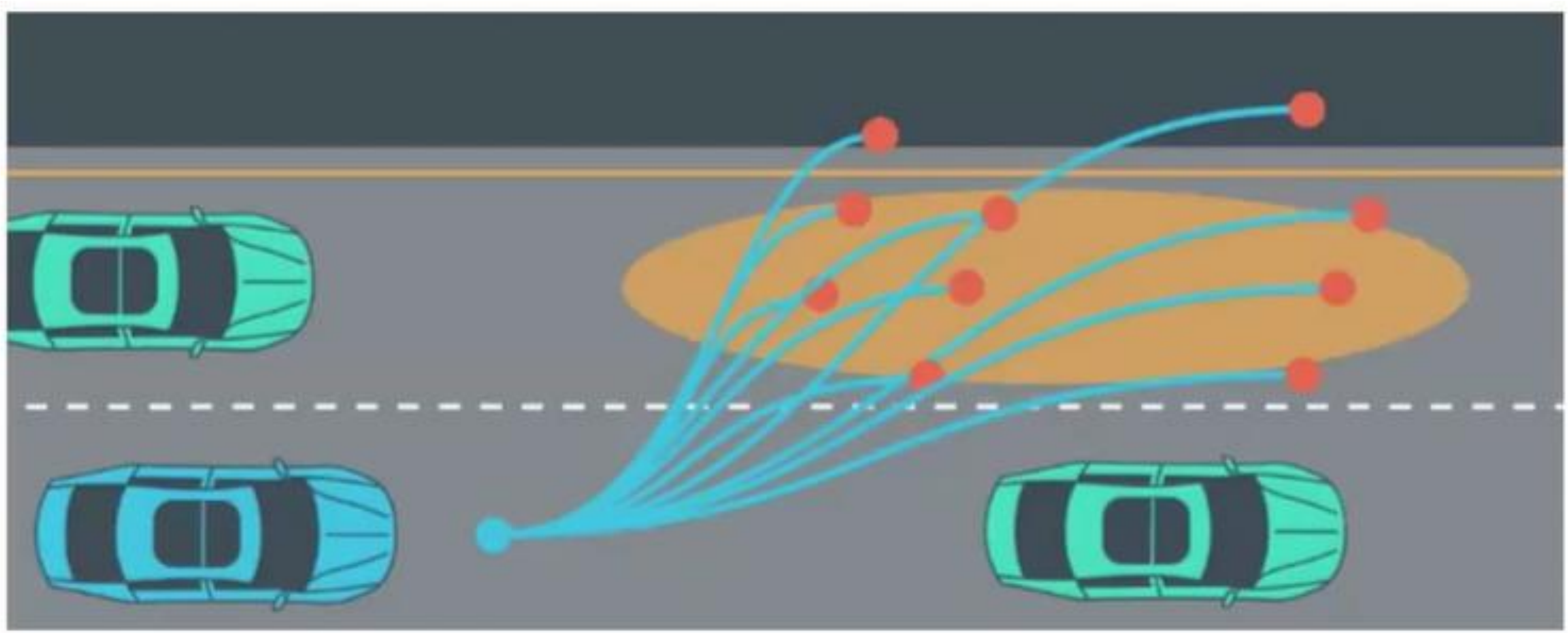
Behavior Planner

- Plan the set of high-level driving actions or maneuvers to safely achieve the driving mission under various driving conditions



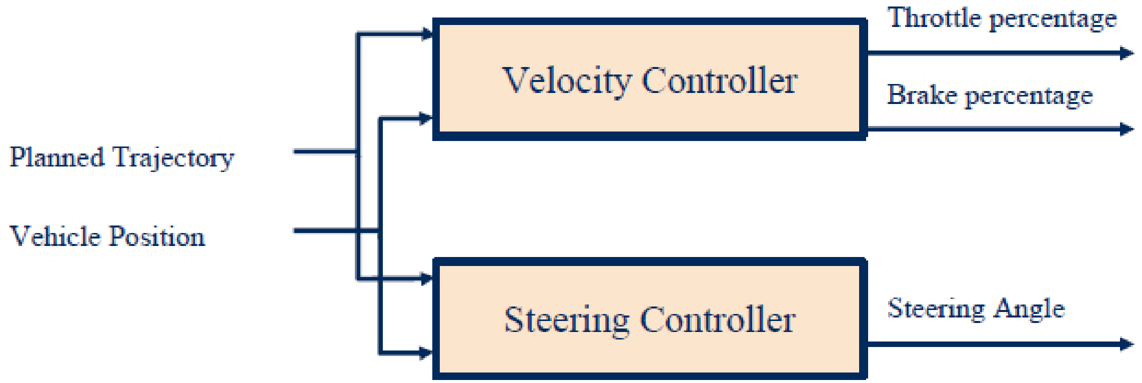
Local Planner

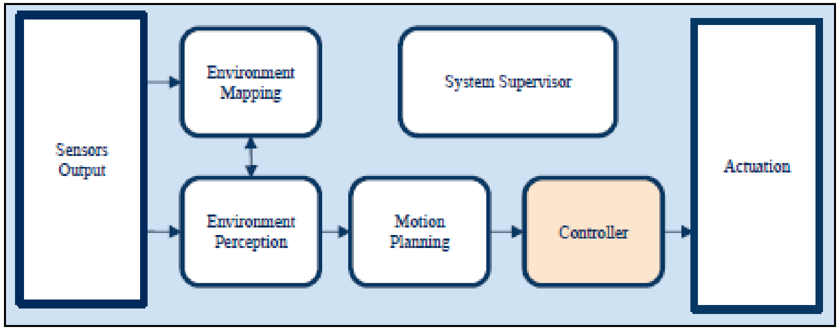
- Plan a safe and smooth trajectory (vehicle pose as function of time)



Controller

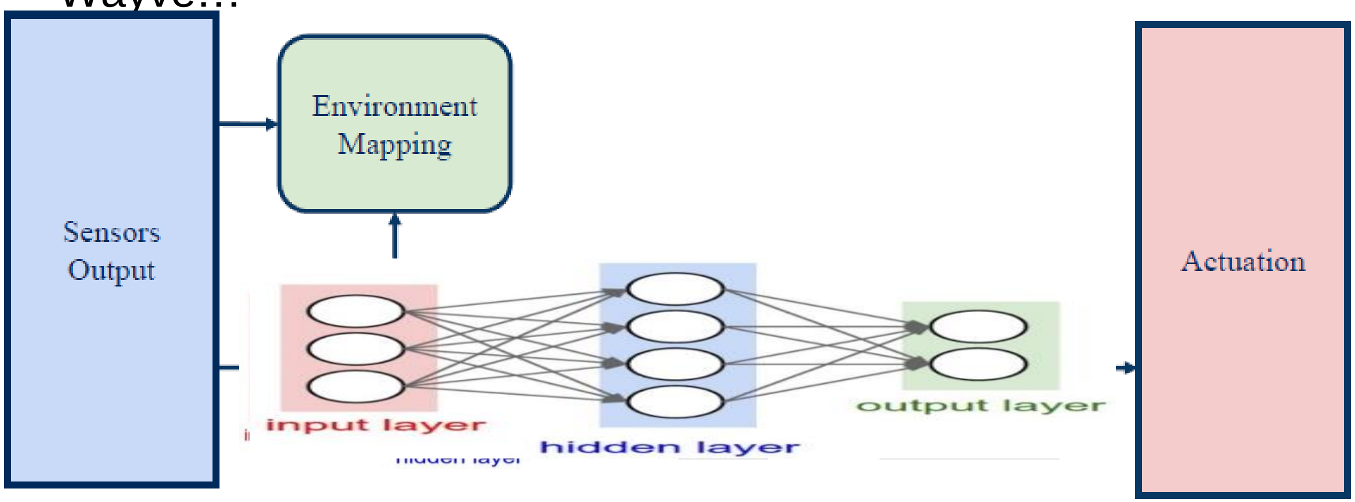
- Velocity controller for longitudinal speed control
- Steering controller for lateral speed control
- Common control algorithms
  - PID: Proportional Integral Derivative
  - MPC: Model-Predictive Control



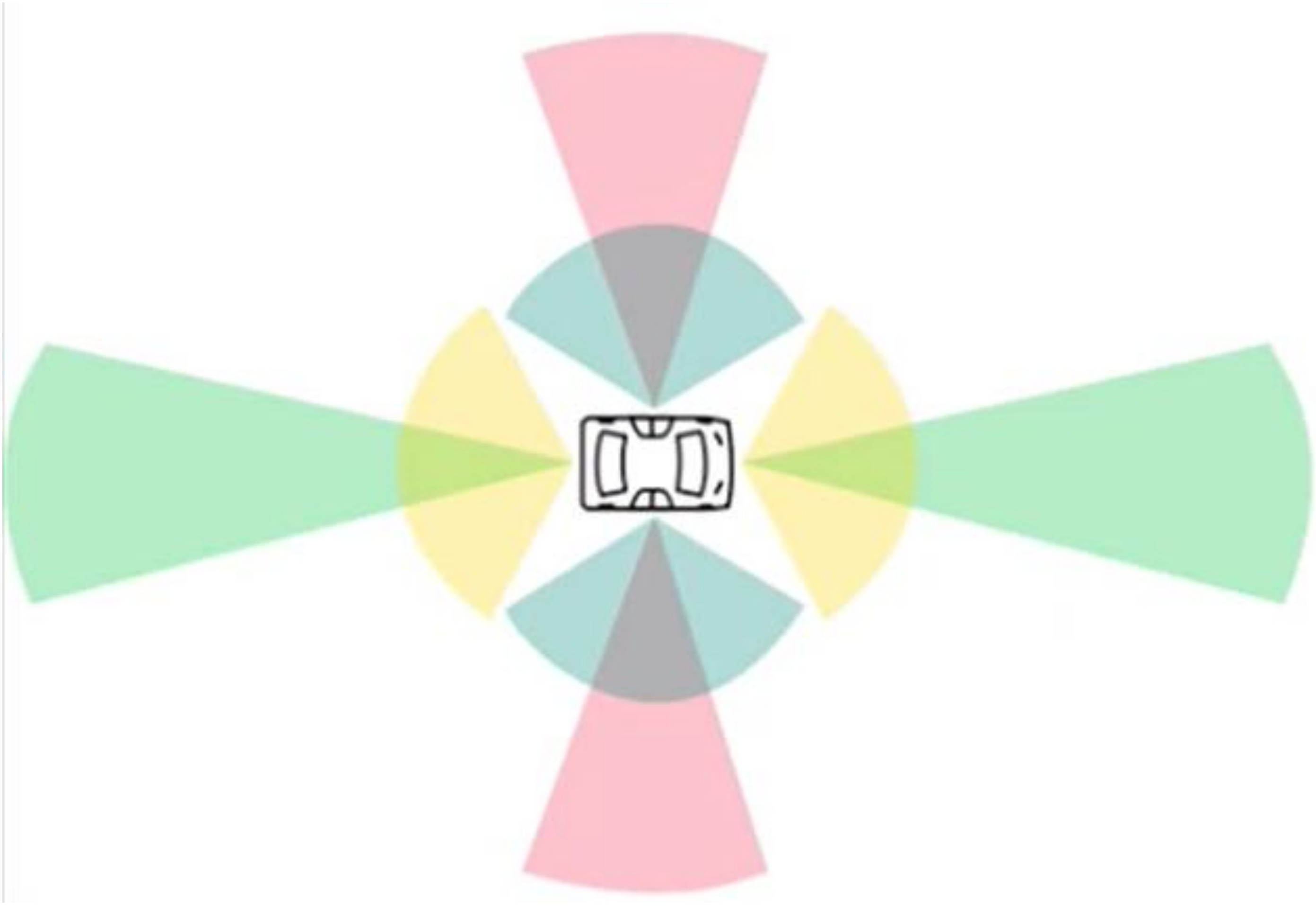


Classic Pipeline vs. ML/DL

- Classic AD processing pipeline: separate algorithms for each processing stage.
- Where does ML/DL come in?
  - CNN (Convolutional Neural Networks) for perception is well accepted.
  - DNN trained with Imitation Learning (IL) or Reinforcement Learning (RL) is still in the early-research stage.
    - “End-to-end” mapping from pixels to control commands
    - Many variants of hybrid approaches, e.g., “half-way” mapping from pixels to waypoints used for planning
    - Several companies are making a bet on it, incl. Waymo, Voyage, Wayve…

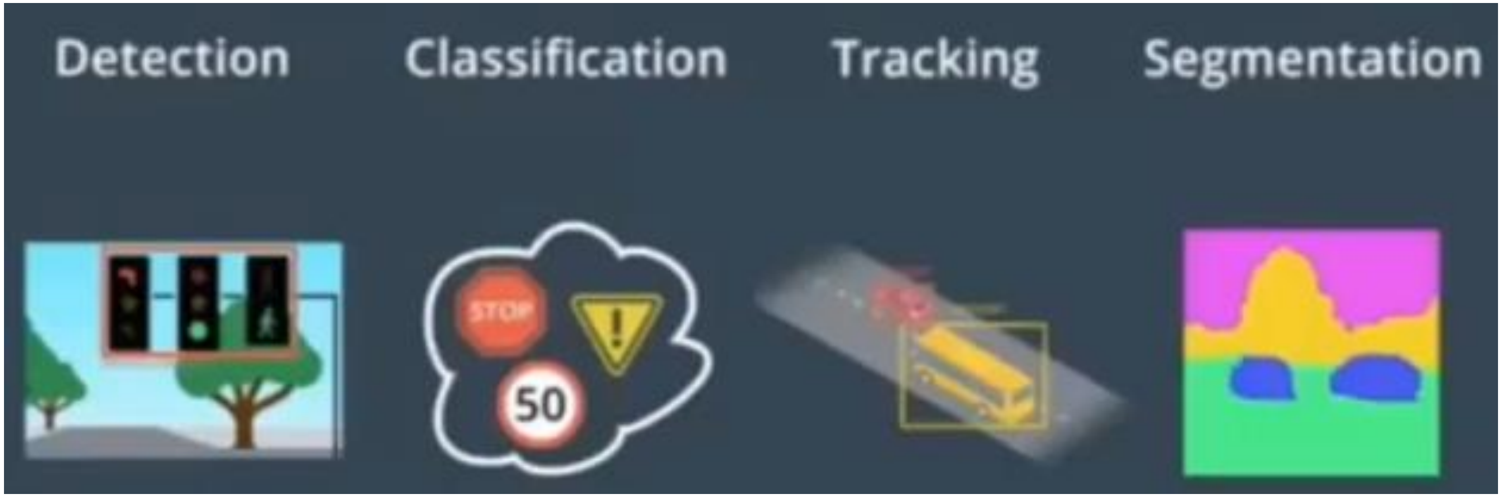


SENSORS AND PERCEPTION

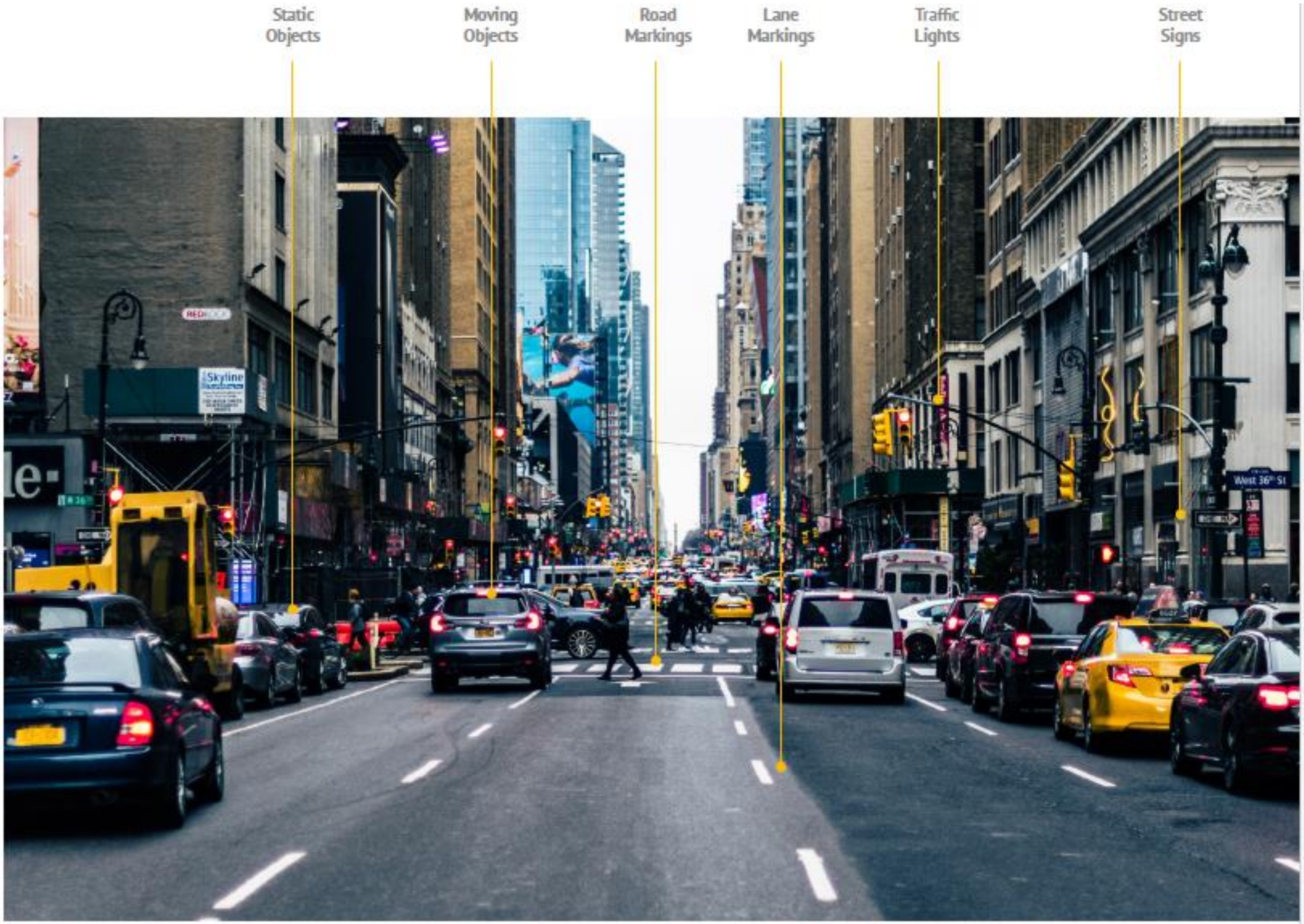


Perception Tasks

- 4 main perception tasks
  - Detection
    - Detect the existence of an object in the environment
  - Classification
    - Identify what the object is, e.g., traffic sign, traffic light, pedestrian
  - Tracking
    - Track a moving object across time
  - Segmentation
    - Semantic segmentation: classify each pixel to its semantic category, e.g., road, car, sky…
    - Instance segmentation: classify each pixel to an object instance, e.g., car1, car2…
- Mobileye's Autonomous Car What the System Sees
  - https://www.youtube.com/watch?v=jKfwHsHUdVc

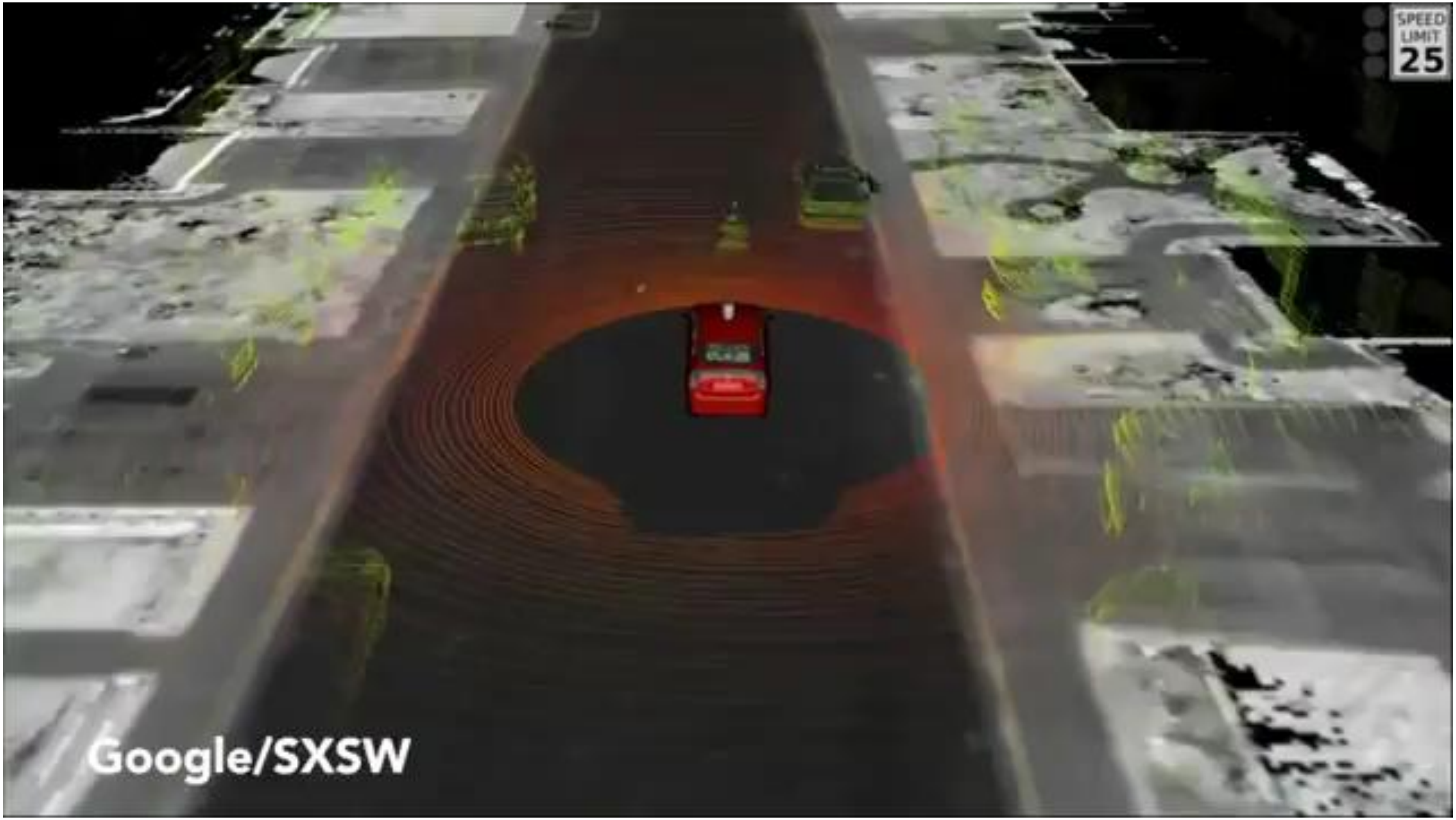


The variety of static and moving objects that an AV needs to detect and recognize

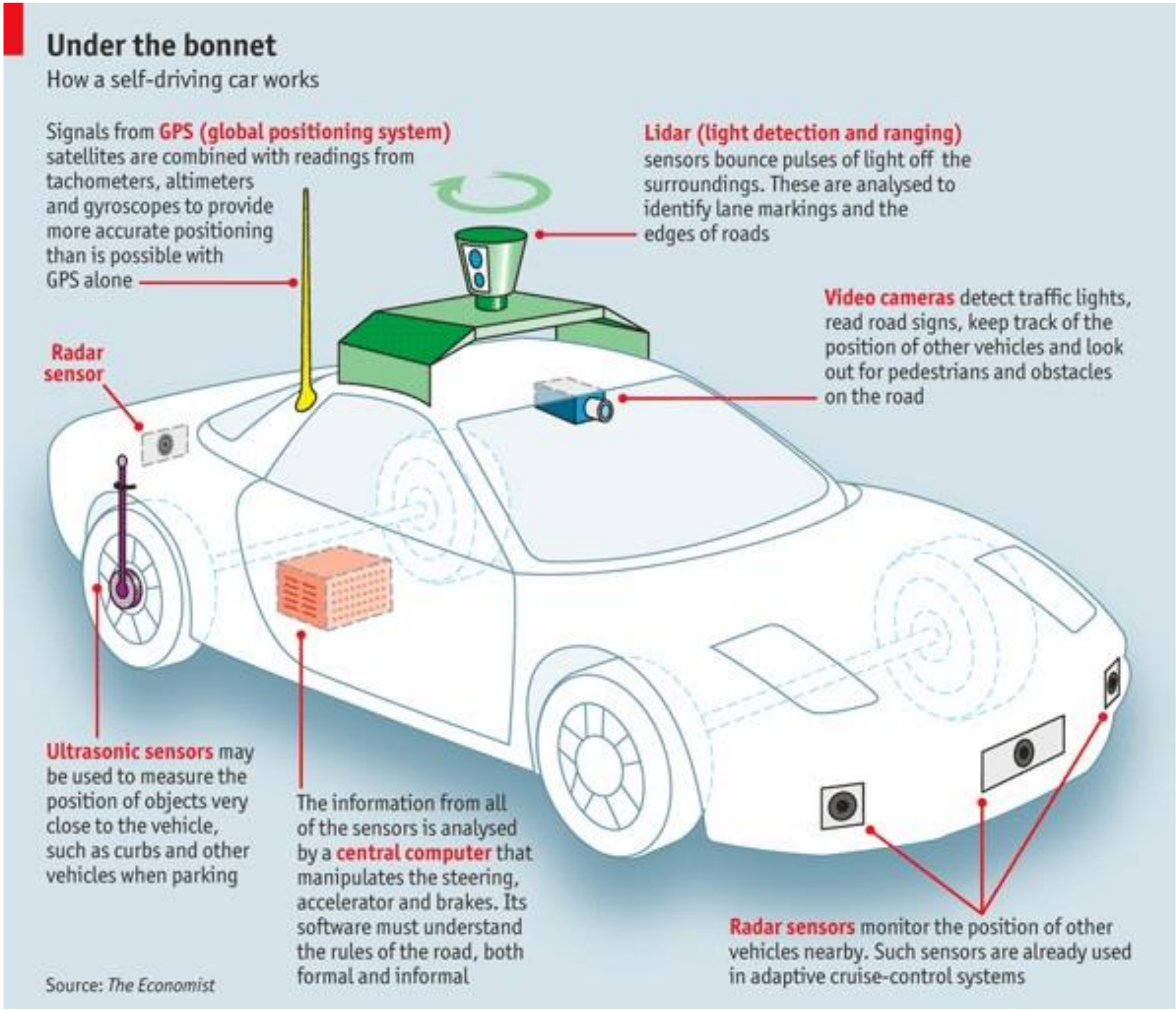


Perception is Challenging

- The long tail distribution is challenging: anything can happen on the road!
- Video from 2015, recorded by Google's AV.



The Typical AV Sensor Configuration



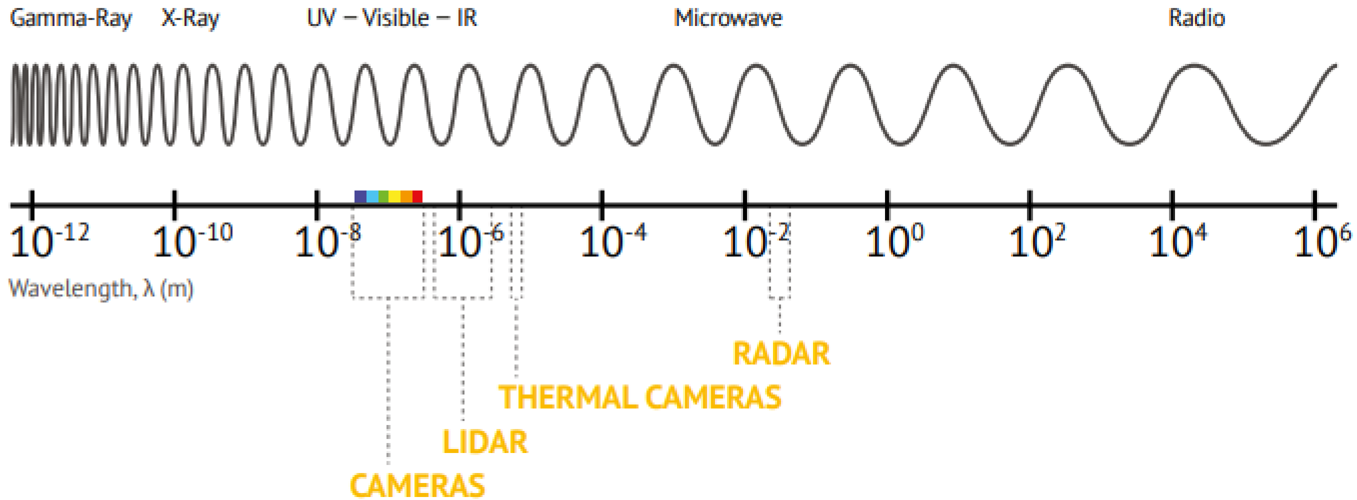
Configuration of Sensors of Some Research AVs

	Uber	Waymo	GM Cruise	Navya Autonomy Cab	Drive.ai	Nissan	Tesla Autopilot V9
Cameras	8	8	16	6	10	12	8
Lidars	1	6	5	10	4	6	0
Radars	4	4	8	4	2	9	1

- Tesla is one of the few AD companies that do not use Lidar.
- Elon Musk, 2017:
  - "Once you solve cameras for vision, autonomy is solved; if you don't solve vision, it's not solved … You can absolutely be superhuman with just cameras."
  - "In my view, Lidar is a crutch that will drive companies to a local maximum that they will find very hard to get out of. Perhaps I am wrong, and I will look like a fool. But I am quite certain that I am not."

Passive vs. Active Sensors

- Passive sensors detect existing energy, like light or radiation, reflecting from objects in the environment.
  - Cameras
- Active sensors (also called range sensors) send their own signal and sense its reflection
  - Lidar, Radar, ultrasound

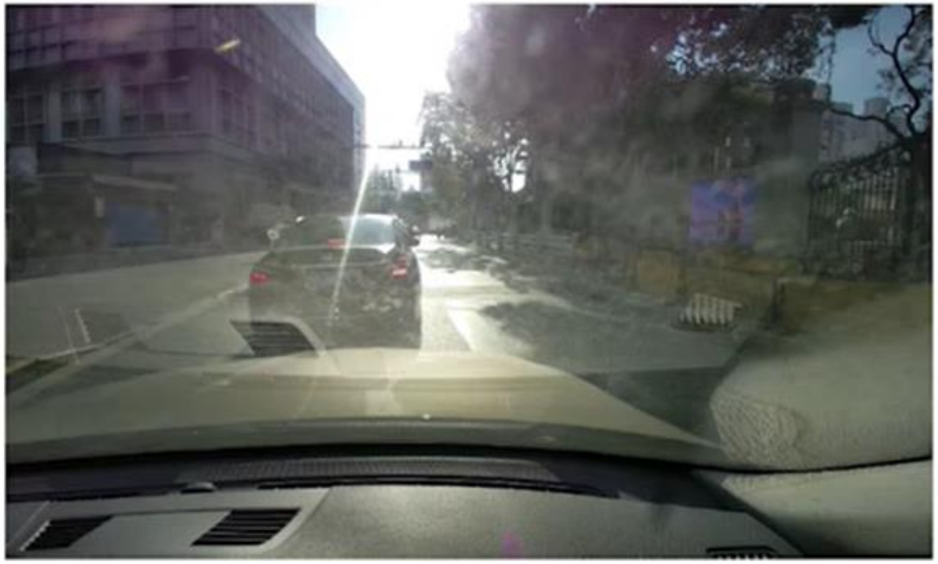


The electromagnetic spectrum and its usage for perception sensors .[16]

https://www.wevolver.com/article/2020.autonomous.vehicle.technology.report

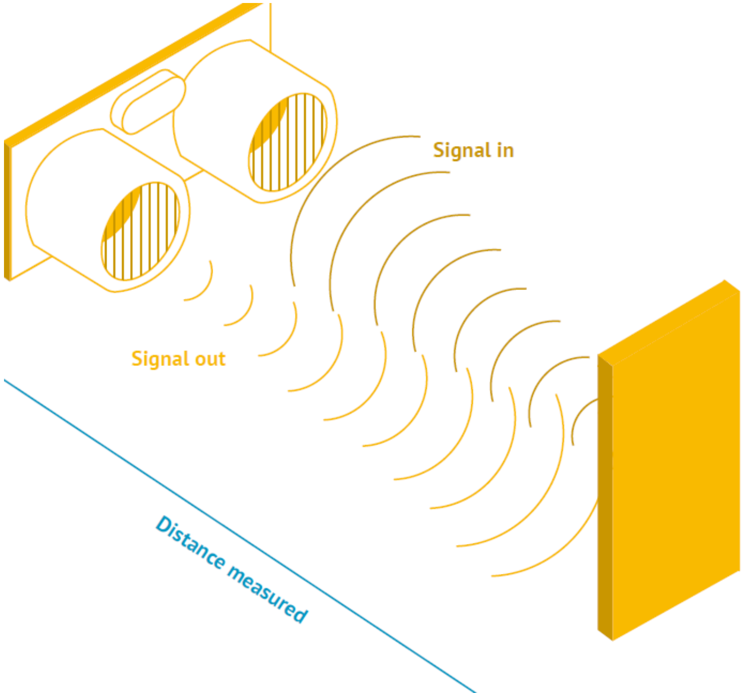
Camera

- Pro: cheap, versatile, stereo vision w. two cameras
- Con: easily affected by illumination conditions, needs additional light at night
- Key parameters
  - Resolution
    - e.g.,1080p HD cameras provide 1920x1080-pixel resolution, or 2.1 megapixels.
  - Field of View (FOV)
    - The extent of the observable world that is seen at any given moment
    - Given same resolution, wider FOV results in large image distortion.
  - Dynamic range
    - Maximum difference between the darkest and lightest pixel intensities in an image, measured in dB. An AV needs HDR (High-Dynamic-Range) cameras with at least 100dB.



Range Sensors

- They rely on Time of Flight (ToF) to measure distance (range), a key element for localization and environment modeling
  - Lidar uses electromagnetic waves.
  - Radar uses radio waves
  - Ultrasonic uses sound waves
- The traveled distance of a wave is given by $d = \frac{v * t}{2}$
  - d: distance
  - v: speed of wave propagation
  - t: ToF (roundtrip)

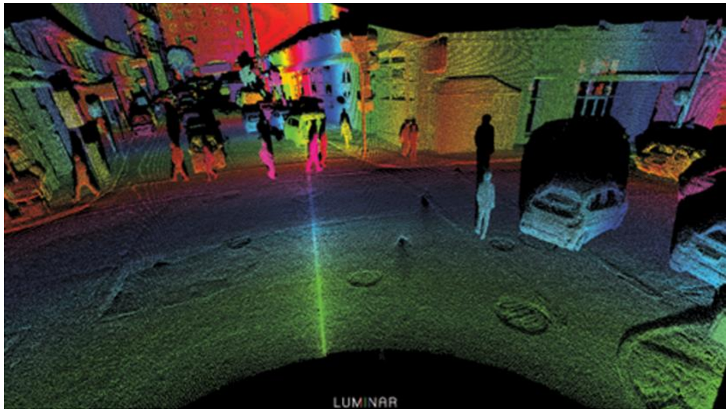


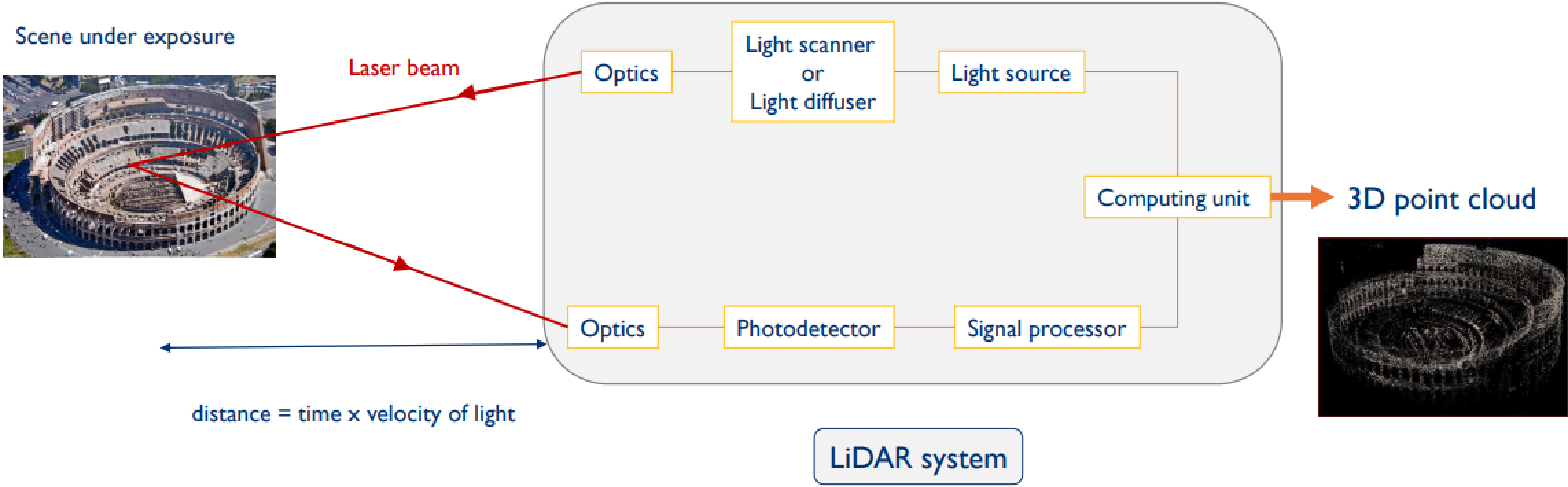
Range Sensors

- Propagation speed v
  - Sound: 0.3 m/ms
  - Electromagnetic wave (incl. light): 0.3 m/ns
    - 1 M times faster than sound
- To travel 3 meters:
  - 10 ms for ultrasonic sensor
  - 10 ns for Lidar
  - Measuring time of flight with electromagnetic signals is not an easy task. Hence Lidars are expensive and delicate
- The quality of range sensors mainly depends on:
  - Inaccuracies in the time of fight measurement (laser range sensors)
  - Opening angle of transmitted beam (especially ultrasonic range sensors)
  - Interaction with the target (surface, specular reflections)
  - Variation of propagation speed (sound)
  - Speed of vehicle and target

Lidar

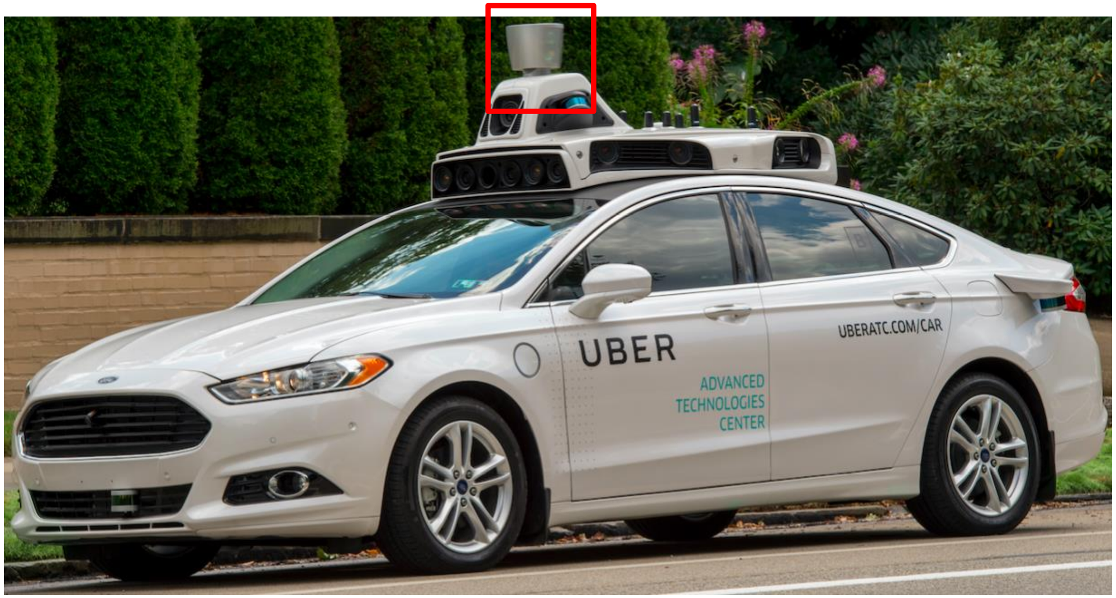
- Lidar (Light Detection and Ranging Device) sends millions of light pulses per second in a well-designed pattern to generate “Point Clouds” that describe the 3D geometry of the surrounding environment
- Pro: independent of lighting conditions, precise distance measurements for 3D perception
- Con: expensive, medium resolution
- Key parameters:
  – Laser beam count
  – Rotation Speed
  – FOV
  – Range distance (from tens to hundreds of meters)



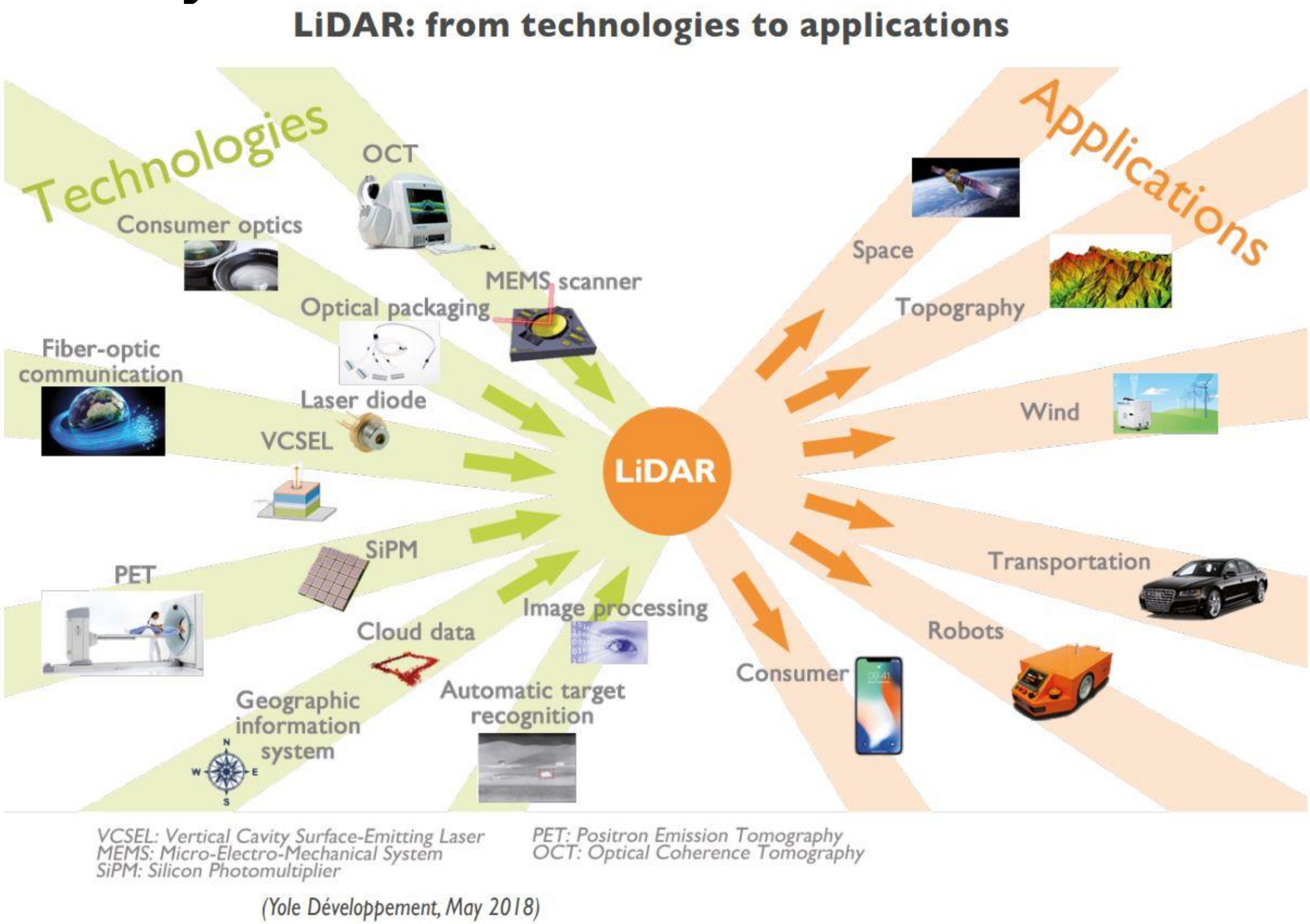


Velodyne Lidar

- The high-end Velodyne HDL-64E with 64 laser emitters
  - Rotation rate up to 15 Hz
  - FOV is 360° horizontally and 26.8° vertically
  - Angular resolution is 0.09° and 0.4° respectively
  - Delivers ~1.3M data points per second
  - Expensive (~USD$40-80K) (cheaper versions available)



Many Variants of Lidars in the Market



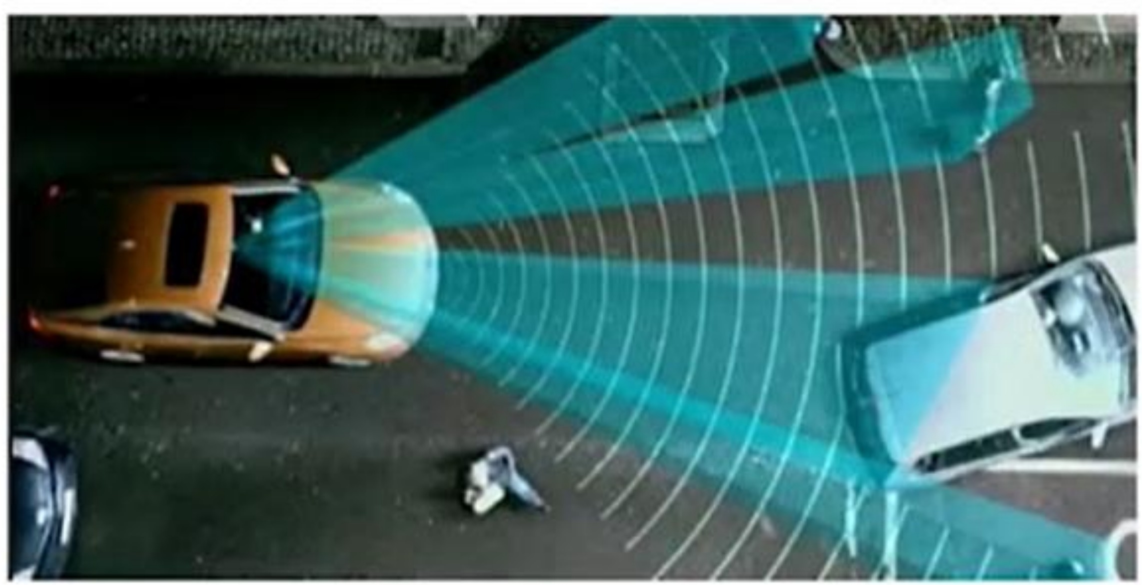
http://www.f4news.com/2018/05/04/yole-on-lidar-market/

Apple iPAD Pro has Built-in Lidar (2020)

- It allows users to scan a depth-accurate depiction of the environment.
- Main application: Augmented Reality (AR)
  - Needs depth information to place virtual objects in the environment.

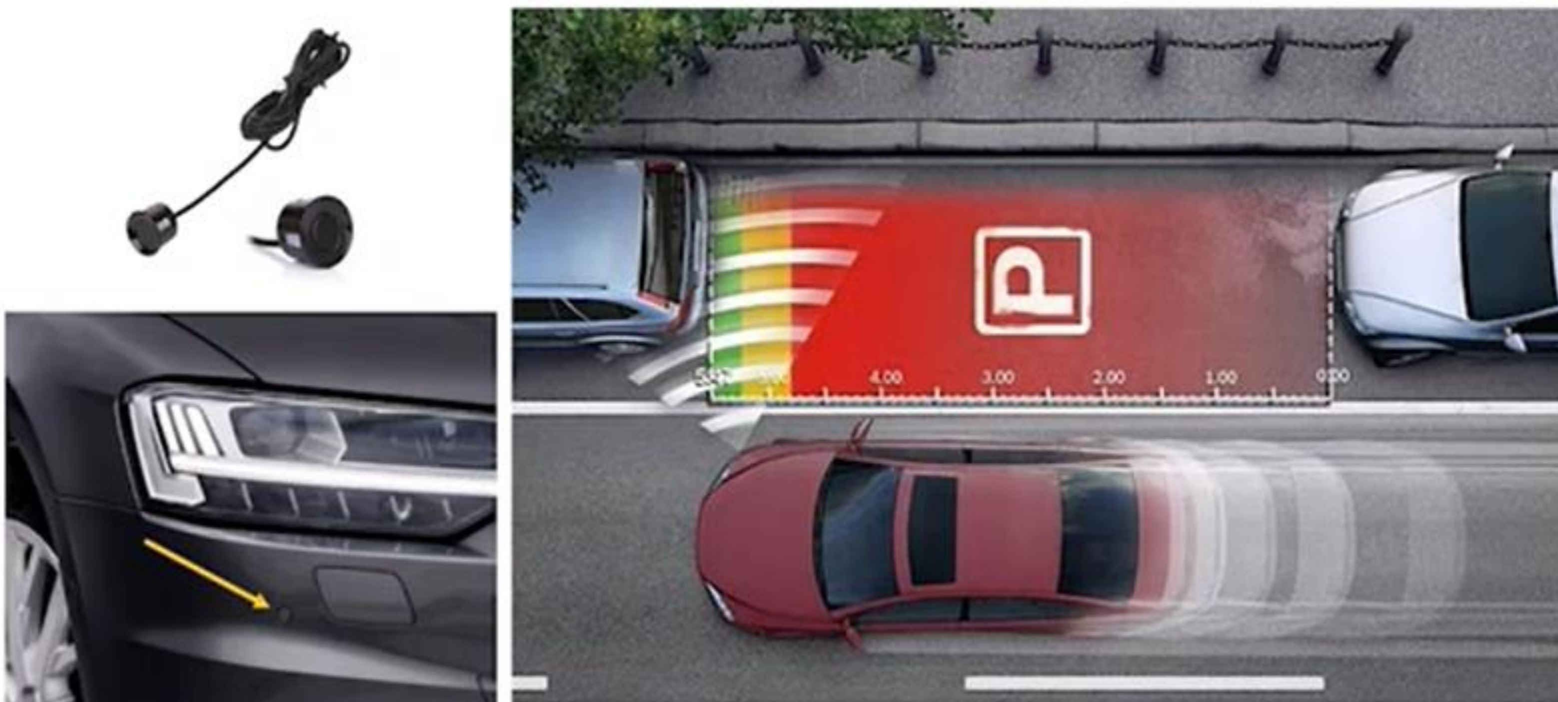
mmWave Radar

- Pro: provides both position and relative speed information; can operate in varied conditions (low-lighting, rain, fog…)
- Con: low resolution
- Key parameters
  - Sensing distance, FOV, Position and velocity accuracy
- Two types
  - Short-medium range with wide FOV
  - Long range with narrow FOV

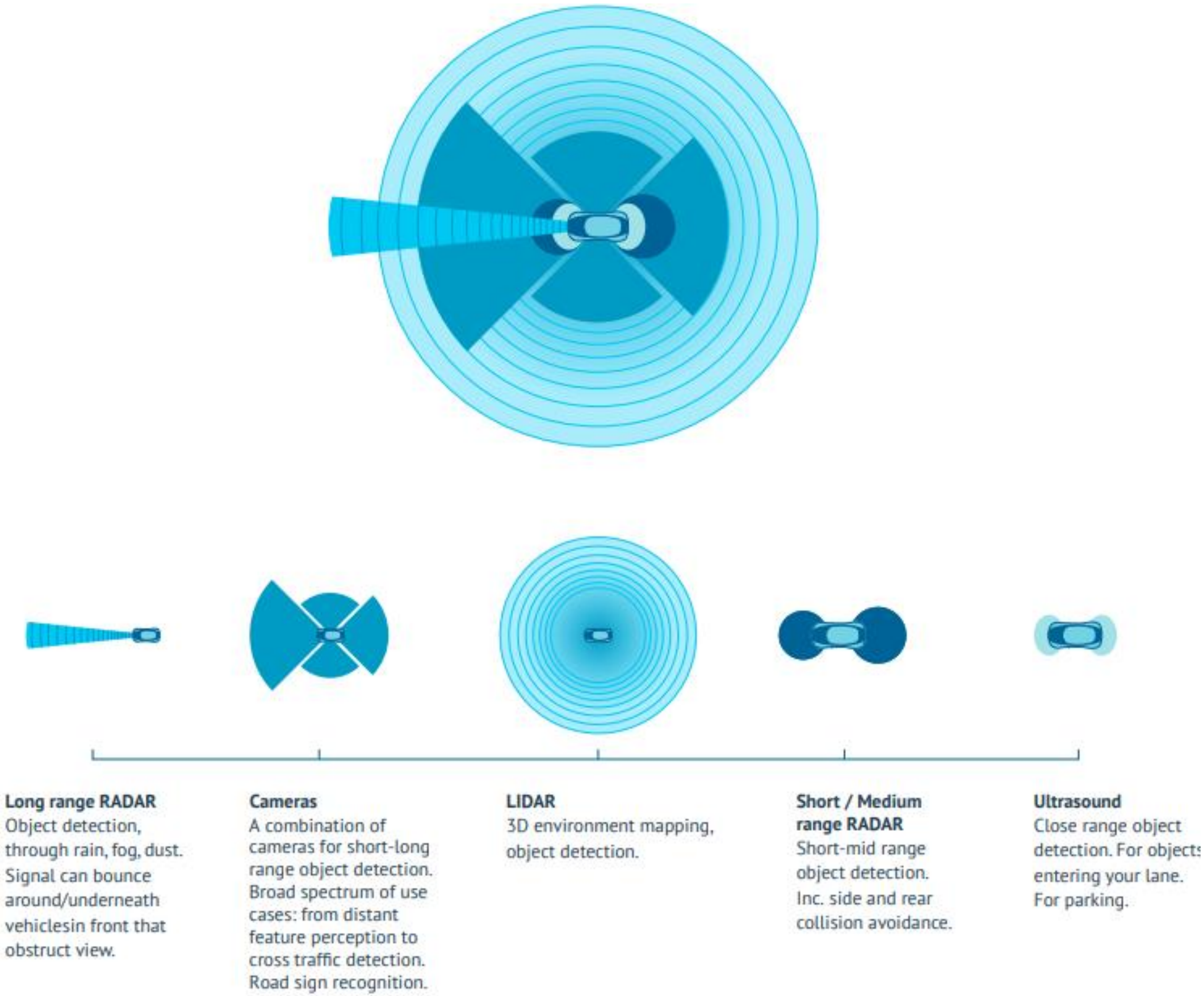


Ultrasound

- Pro: not affected by lighting conditions, rain or fog
- Con: short sensing range (mainly used for parking assistance)
- Key parameters
  - Sensing range
  - FOV



Comparison of Sensing Ranges



https://www.wevolver.com/article/2020.autonomous.vehicle.technology.report

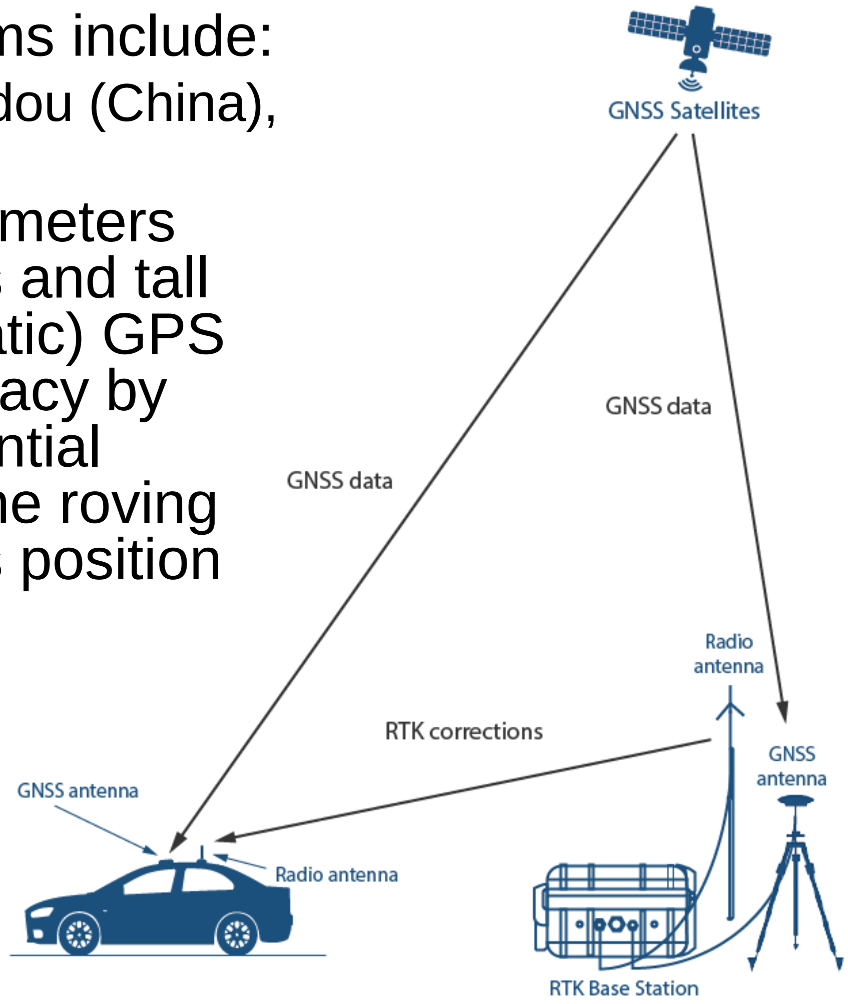
Other Comparisons

- Each sensor has its strengths and weaknesses
- Sensor fusion crucial for robust perception

	Camera	**Lidar**	**Radar**
Sensing Range	Mixed	Mixed	Good
Functioning in bad weather	Poor	Mixed	Good
Functioning in poor lighting	Mixed	Good	Good
Object Detection	Mixed	Good	Mixed
Object Classification	Good	Mixed	Poor
Lane Tracking	Good	Poor	Poor

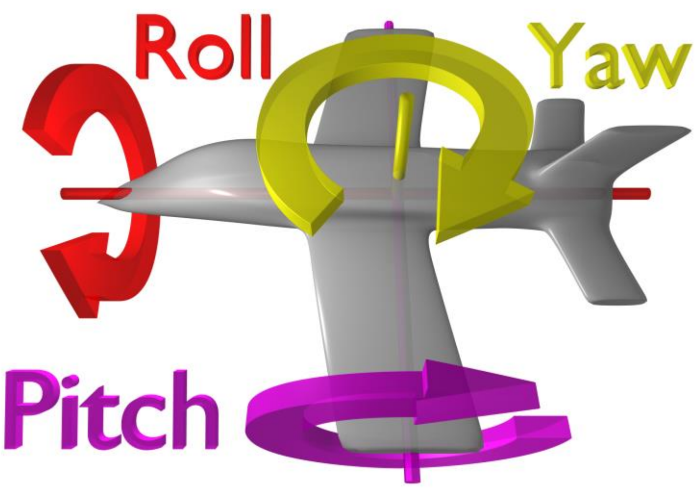
GNSS/GPS

- GNSS (Global Navigation Satellite System) provides localization service for outdoor applications. Current GNSS systems include:
  - GPS (USA), Galileo (Europe), Beidou (China), GLONASS (Russia)
- Conventional GPS provides a few meters accuracy, affected by cloud covers and tall buildings; RTK (Real-Time Kinematic) GPS can provide centimeter-level accuracy by calculating and transmitting differential correction data via radio to allow the roving GPS system (vehicle) to correct its position



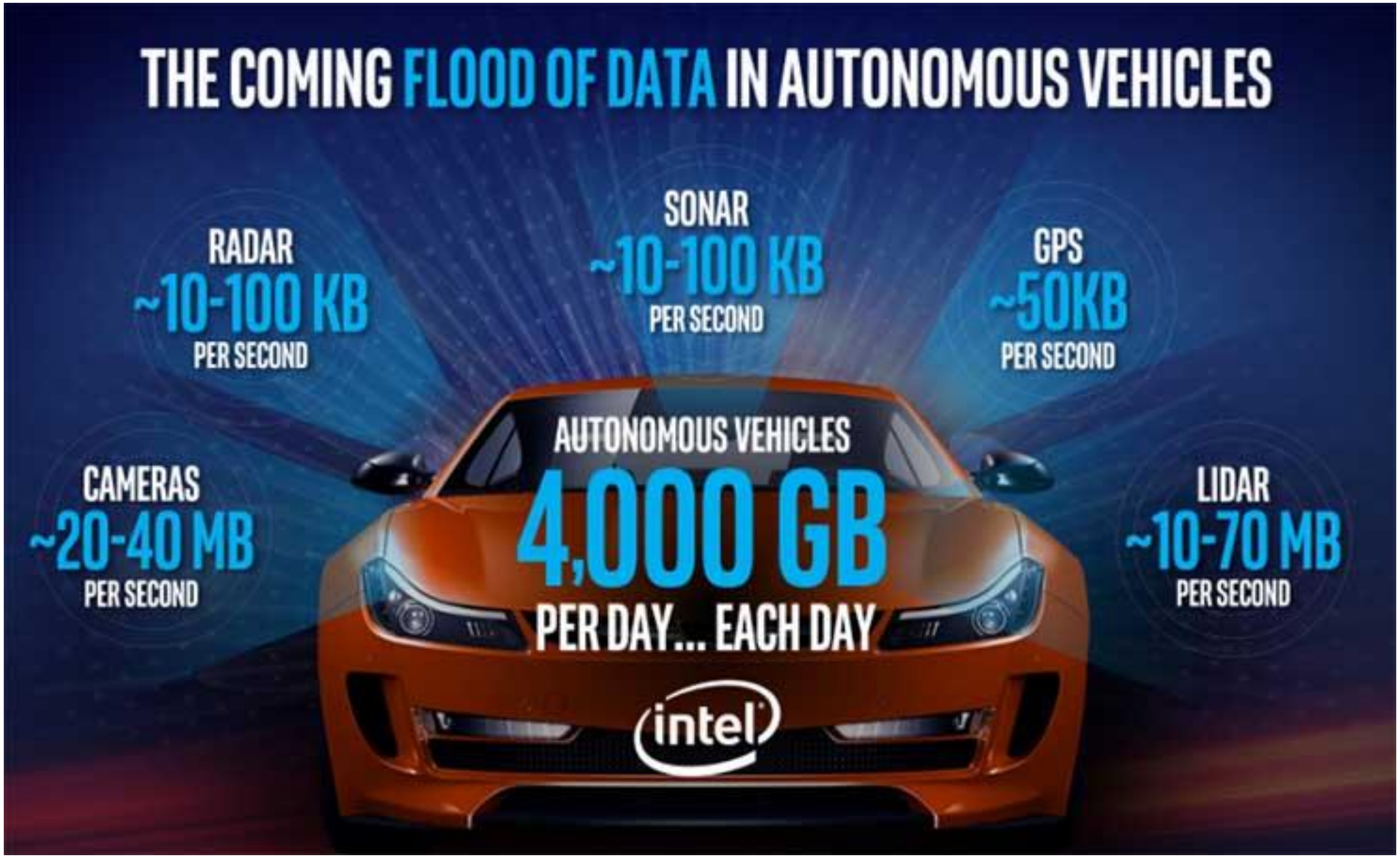
IMU

- IMU (Inertial Measurement System) measures acceleration (linear and angular) and orientation (yaw, pitch, roll)
  - For ground vehicles on 2D plane, only yaw is relevant.
  - For aerial vehicles in 3D space, all three are relevant
- Often combined with GNSS/GPS to form INS (Integrated Navigation System), using sensor fusion to achieve higher estimation accuracy



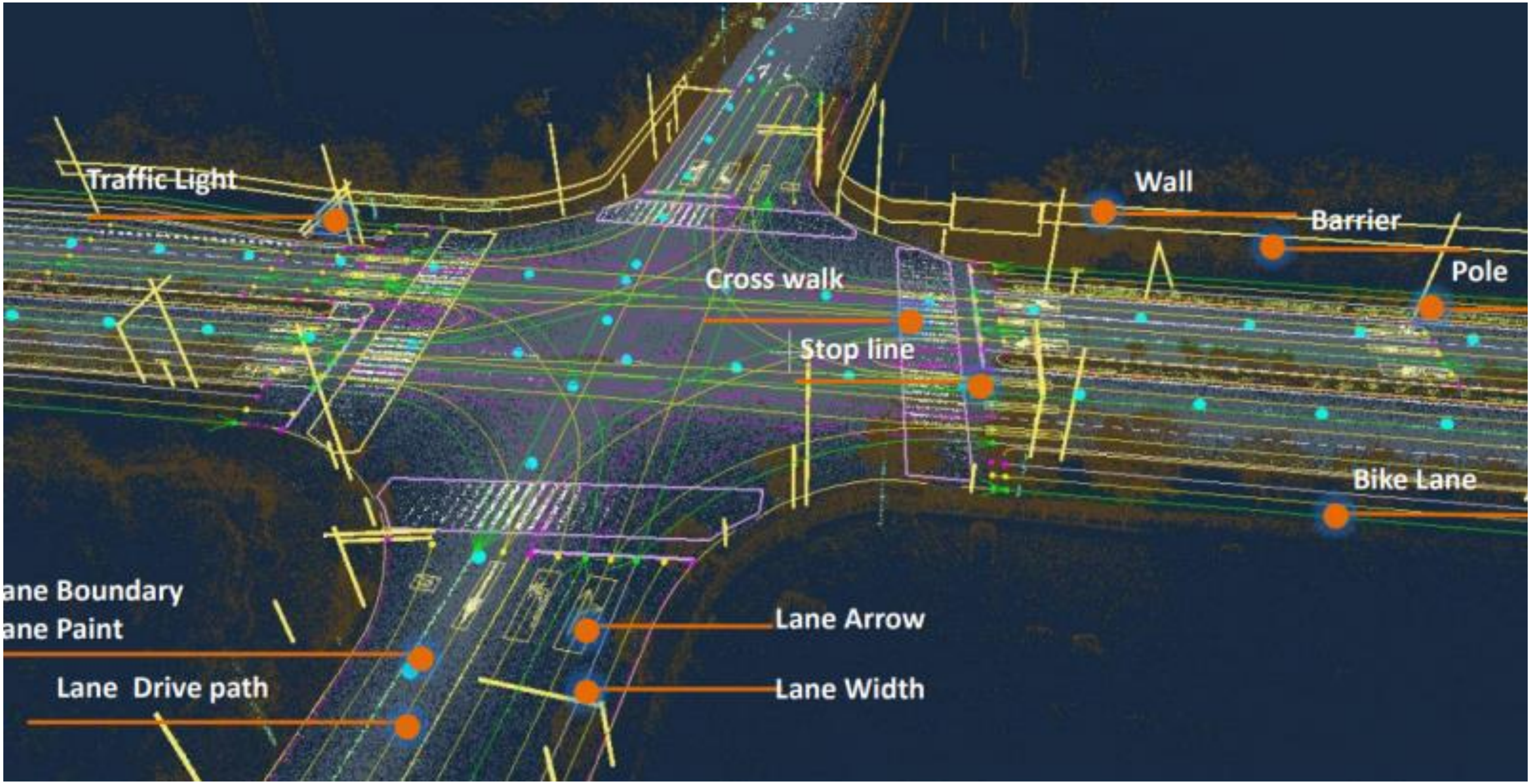
AV Sensors Generates Big Data

- Sensors, esp. cameras and lidars, generate the most amount of data
- Sensor data must be processed in real-time by perception algorithms



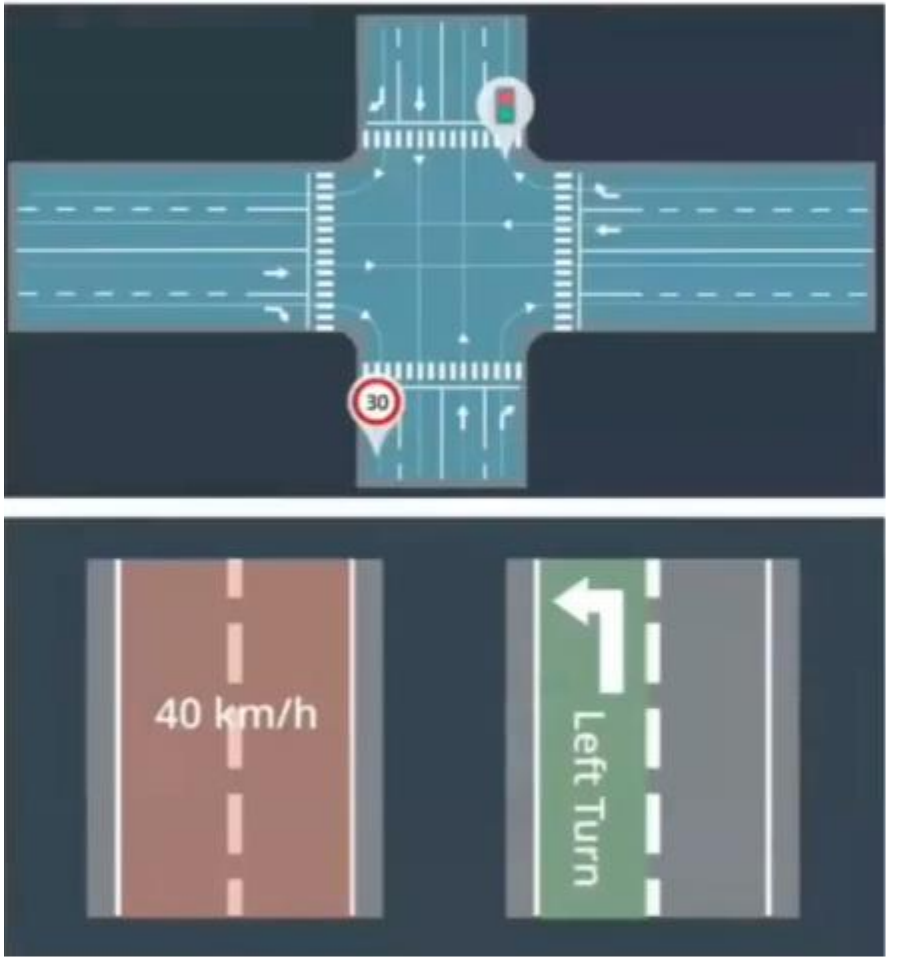
https://datacenterfrontier.com/autonomous-cars-could-drive-a-deluge-of-data-center-demand/

High-Definition (HD) Maps



HD Maps

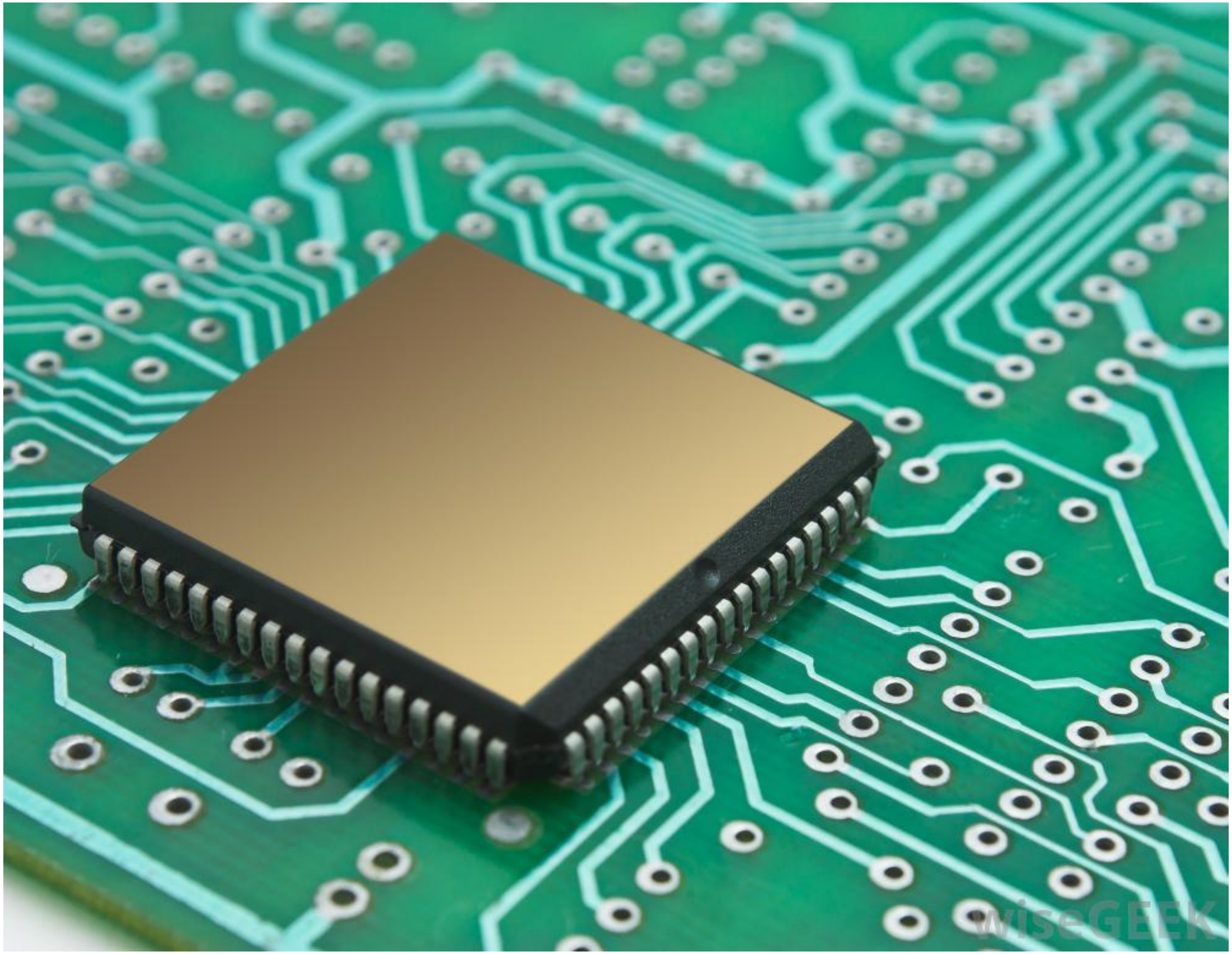
- Different from navigation maps (e.g., Google Maps) designed for human eyes, HD maps are designed for processing by computers
  - Highly-accurate (centimeter-level) 3D representation of the road network, e.g., cross section layout; locations of traffic lights/signs; semantic information on certain road segments (speed limits…)
- Benefits
  - Help reduce Region-of-Interest (ROI) for detection of traffic elements
  - Help with AV localization based on known object positions
  - help with recognition of lane center line



Are HD Maps Necessary?

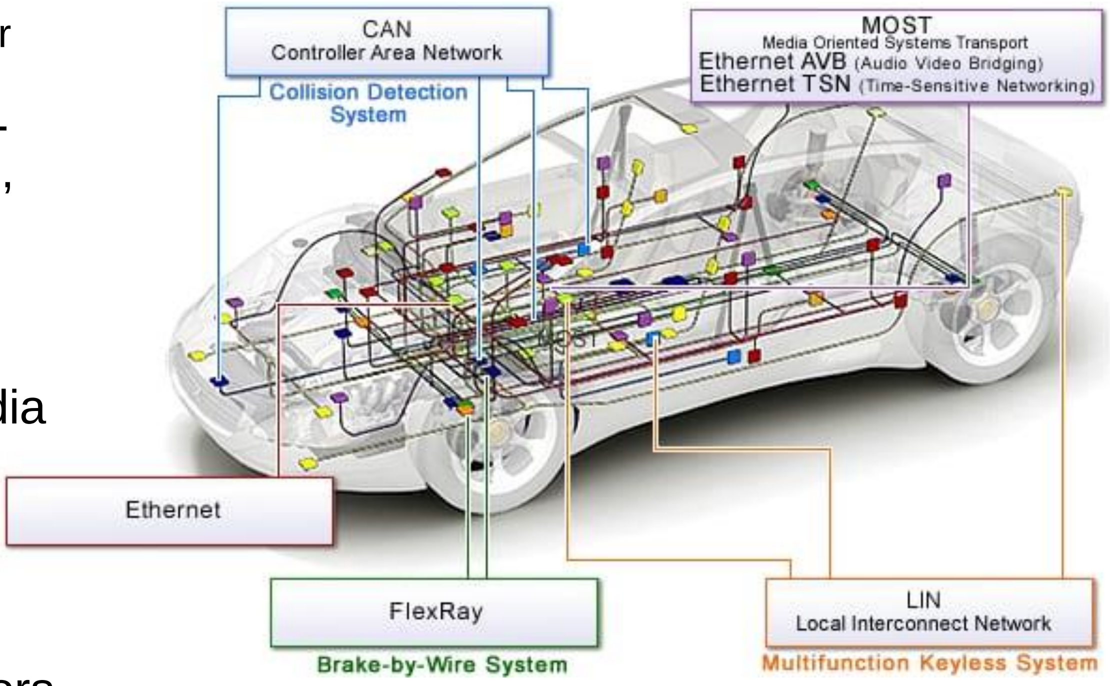
- Humans don’t need HD maps to drive
- HD maps vs. on-board sensing
  - Use of HD maps limits the area of operation; mapless systems allow universal operation anywhere.
  - Use of HD maps reduces the computational burden on on-board sensing: the more that has already been mapped out, the easier it is for the on-board system to focus on the moving parts; mapless systems must figure out everything on-the-fly with no prior knowledge of the environment
- It is generally agreed that L4-L5 levels of automation cannot work without HD maps, at least for now; L2-L3 levels may work without them (e.g., Tesla does not use HD maps)

HARDWARE PLATFORMS



Typical Automotive E/E Architecture

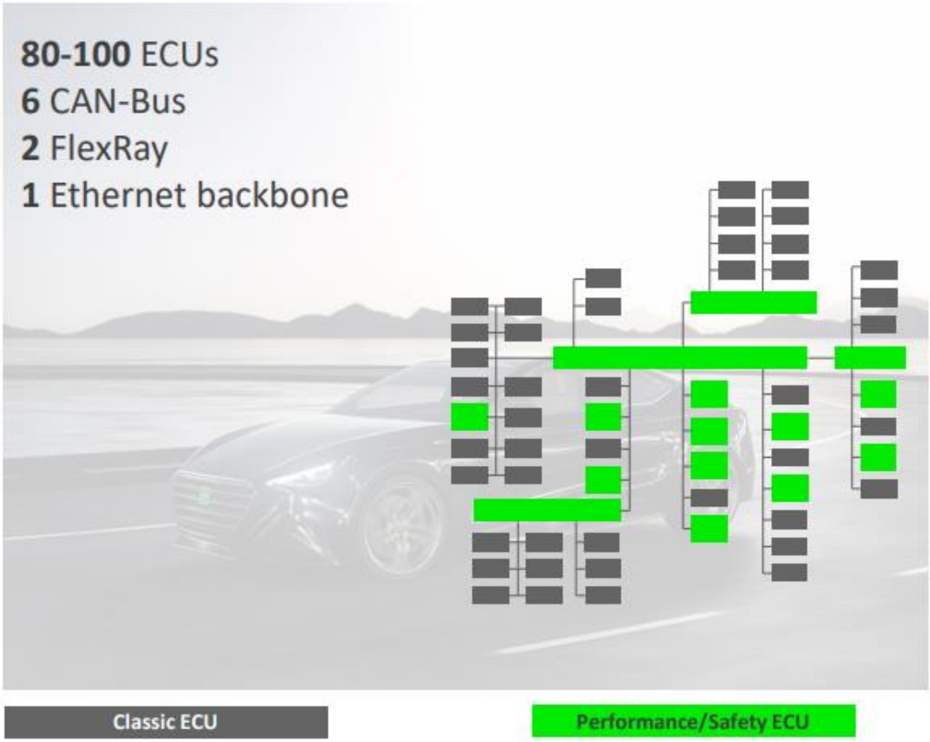
- Ethernet as high-bandwidth backbone network
  - Ethernet TSN posed to be the dominating standard protocol.
  - Regular Ethernet is also used for diagnostics
- FlexRay for safety-critical X-by-Wire, where X stands for brake, steer, drive…
  - Gradually being replaced by Ethernet
- Media Oriented Systems Transport (MOST) for multimedia transmission
  - Gradually being replaced by Ethernet
- CAN (Controller Area Network) for low-bandwidth network and interfacing with sensors/actuators
- LIN (Local Interconnect Network) for body electronics, e.g., door, light, rearview mirrors…

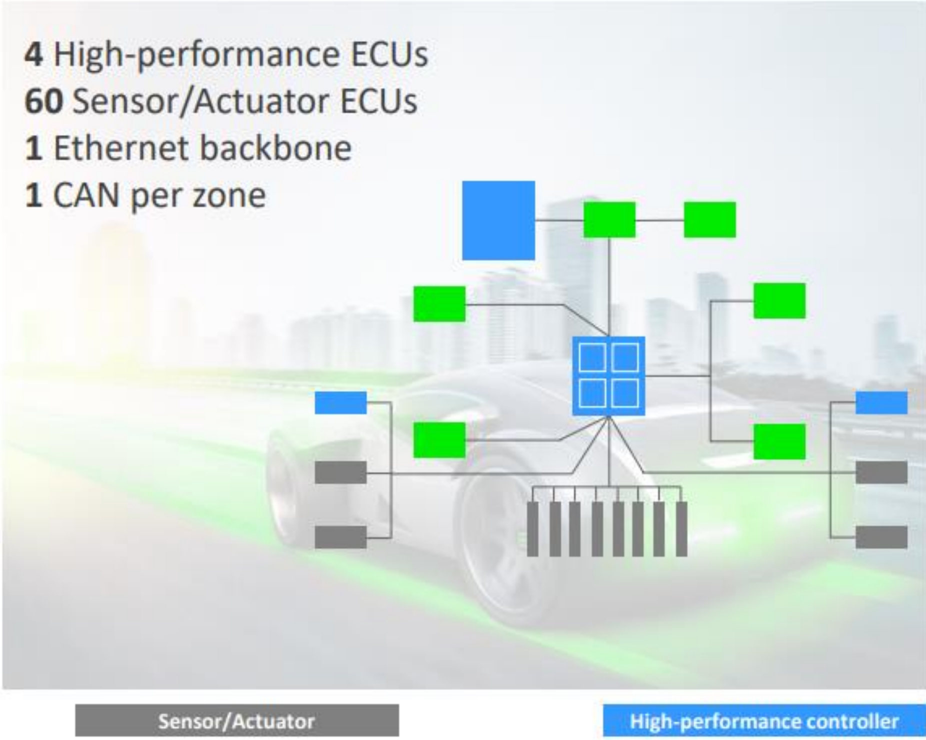


https://www.renesas.com/eu/en/solutions/automotive/technology/networking-solutions.html

Evolution of Automotive E/E Architecture

- From many (~80-100) distributed and networked ECUs to a few (~4) high-performance ECUs with massive computing power, and large number of (~60) small ECUs for interfacing with sensors and actuators.
- This helps simplify system architecture, reduce network load, and improve system reliability.





Trunk of an Experimental AV from Ford (2017)

Where do I put my groceries?

https://www.theverge.com/2017/10/10/16449416/nvidia-pegasus-self-driving-car-ai-robotaxi

AD Hardware Considerations

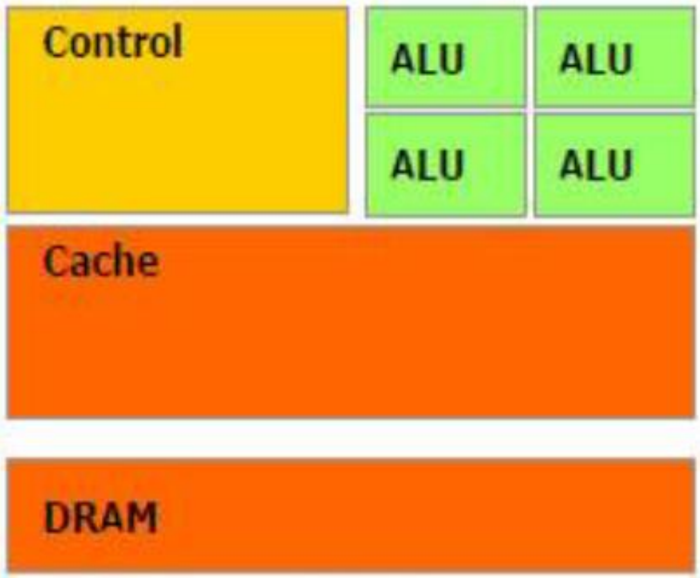
- Power consumption
  - Power consumption of electronics (sensors and computing hardware) for AD may be 100x that of a vehicle with regular ADAS. This drains battery and implies increased fuel consumption or reduced range for EVs
    - EV drivers often turn off air conditioning due to range anxiety; will they turn off AD and drive manually for this reason?
  - Waymo and Ford now focus on Hybrid Vehicles, while Uber uses a fleet of full gasoline SUVs.
- Cooling capacity
  - Fan or liquid cooling.
- Form factor
  - Must be compact and unobtrusive.
- Cost
  - Important for mass deployment.
  - Cost of electronics in an experimental AV often exceeds cost of the original vehicle.

SoC Hardware for AVs

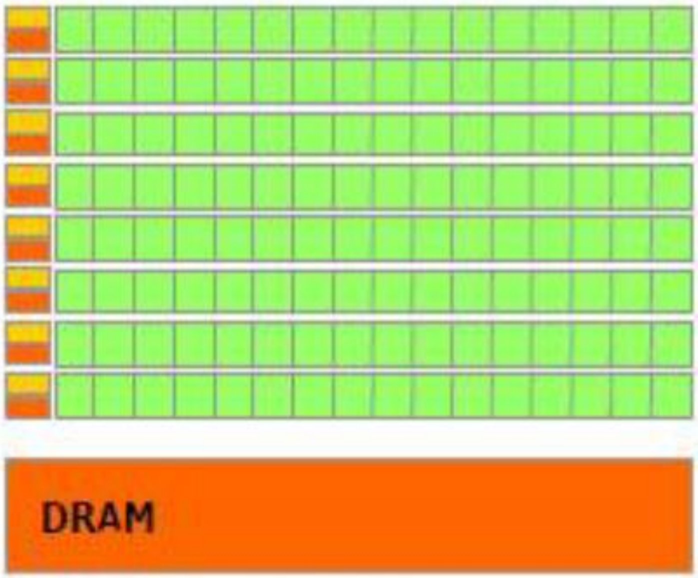
- The most compute-intensive workload is Deep Learning
  - Mostly inference tasks, but may also perform training tasks in case of online-learning.
- Many vendors provide SoC (System-on-Chip) products that integrate CPU cores with specialized computational engines for Deep Learning:
  - GPU (Graphics Processing Unit)
    - NVIDIA is the only serious player.
    - Other GPU venders, e.g., AMD, ARM, Intel, focus on computer graphics instead of general-purpose computing (GPGPU).
  - FPGA (Field-Programmable Gate Arrays)
    - Xilinx, Intel Altera
  - ASIC (Application-Specific Integrated Circuit)
    - An explosion of specialized ASICs for Deep Learning in recent years, with hundreds of companies and products ranging from high-performance to embedded.
  - DSP (Digital Signal Processor)
    - Mainly for image preprocessing, e.g., products from Texas Instruments.

CPU vs. GPU

- GPU has much simpler control logic than CPU, hence has more computational elements (Arithmetic Logic Units)
- GPU is ideally suited for processing highly-parallel workloads
  - e.g. matrix-multiply, which is a core operation in Deep Learning algorithms



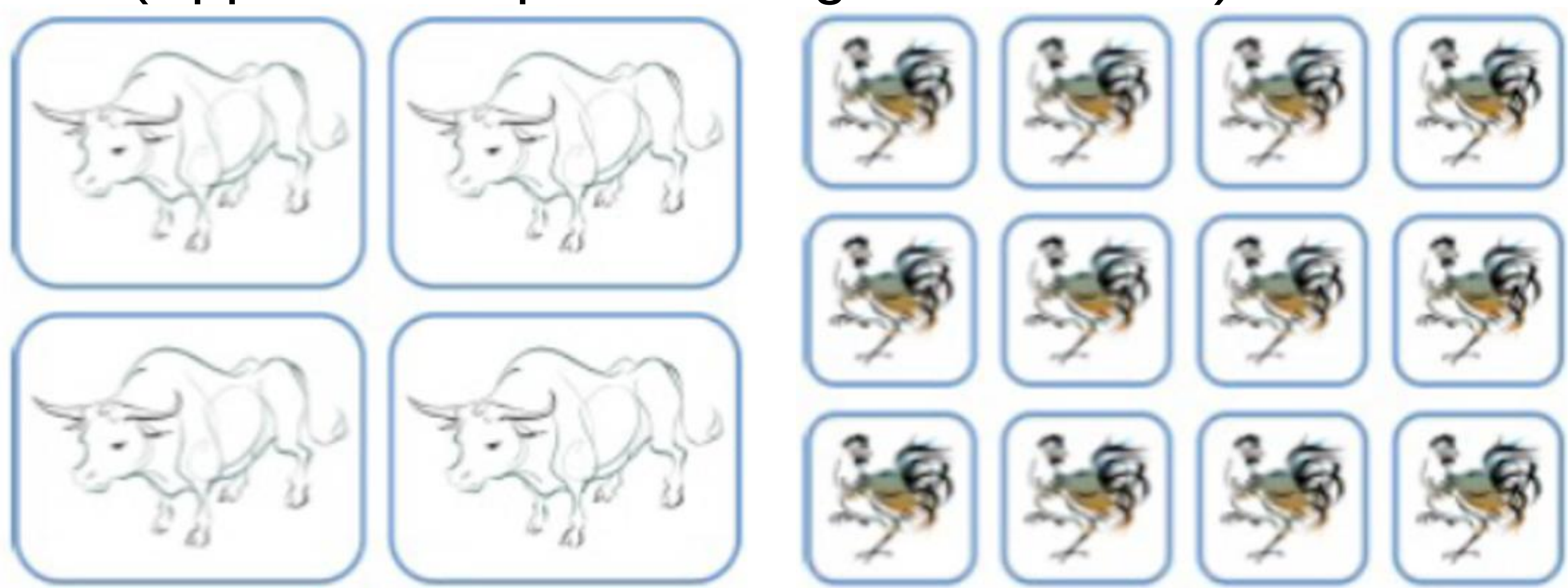
CPU



GPU

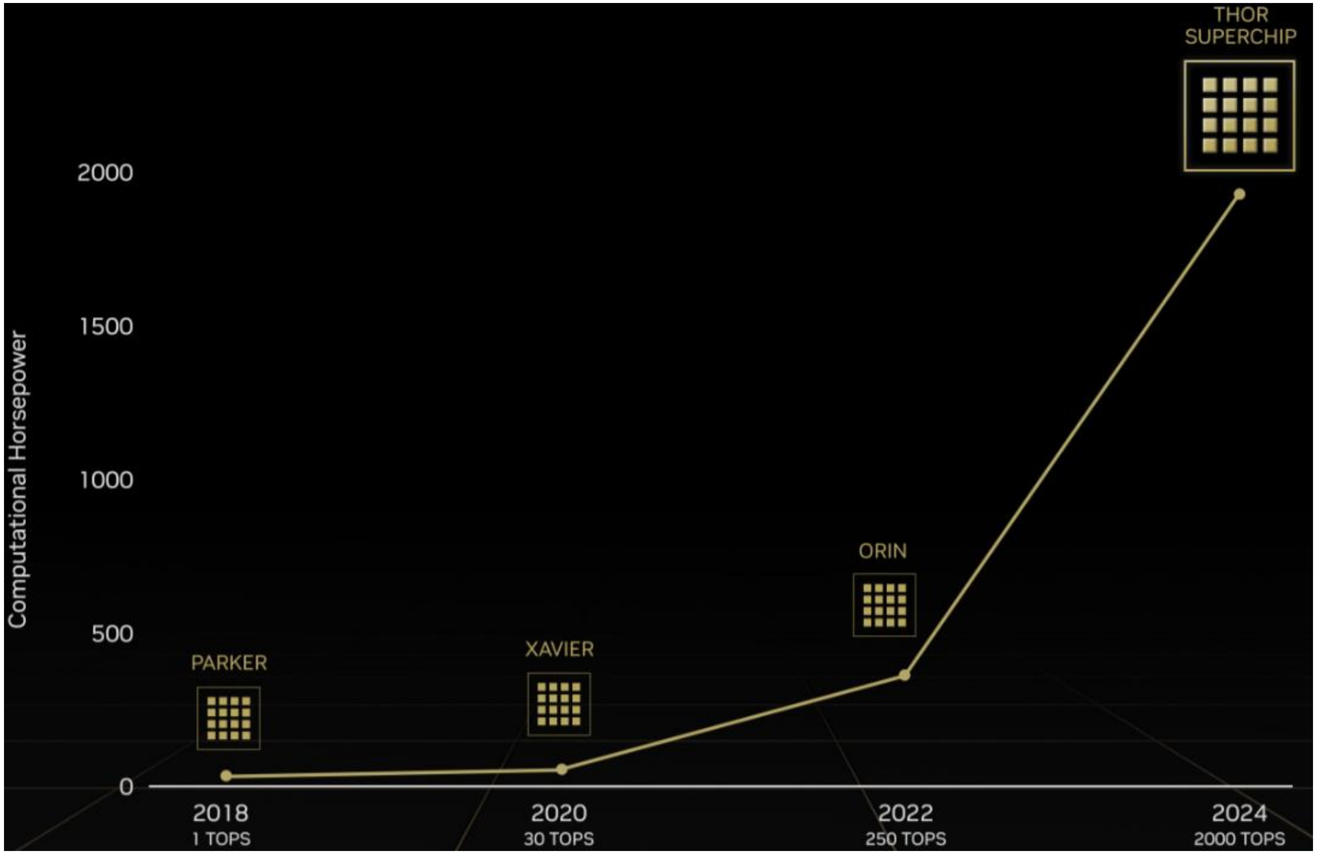
CPU vs. GPU

- When you are plowing a field, would you prefer 4 strong oxen (a multicore CPU), or 1024 chickens (a GPU)?
- Similar arguments for FPGAs (Field-Programmable Gate Arrays) and ASICs (Application-Specific Integrated Circuits)



NVIDIA DRIVE Hardware

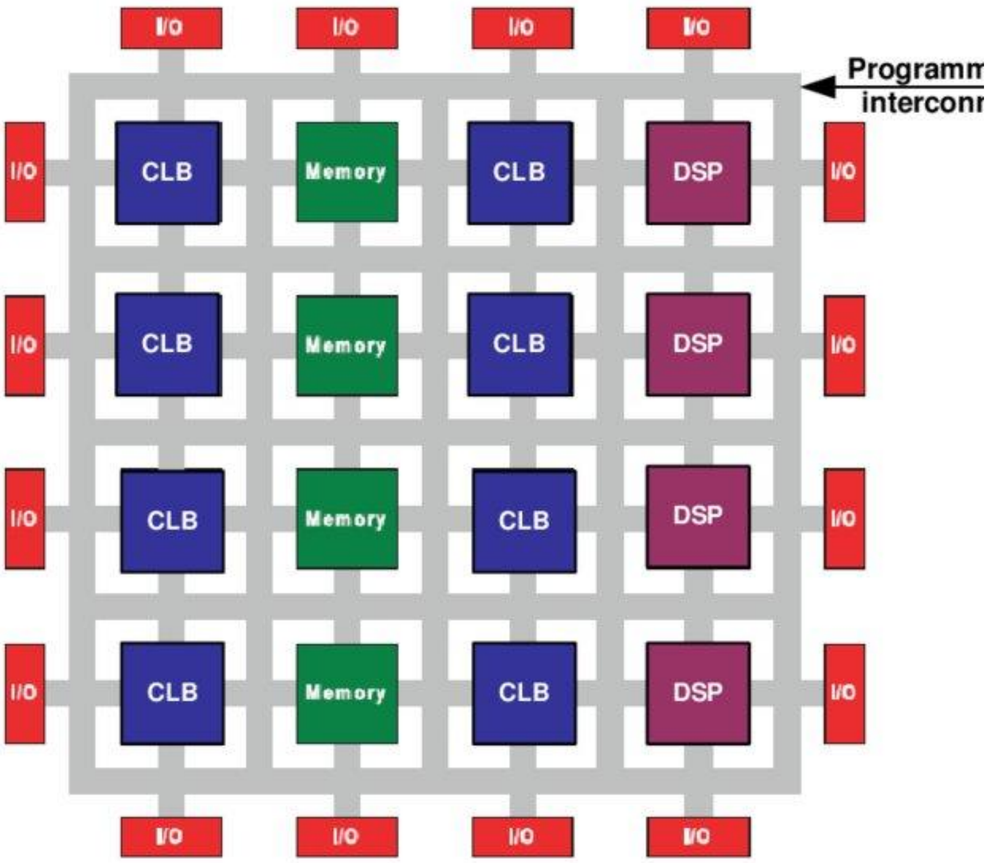
- A family of products, ranging from the low-end Parker to the latest high-end THOR with 2000 TOPS (Tera Operations Per Second)
  - Besides CPU and GPU cores, also includes NVDLA (NVIDIA Deep Learning Accelerator), an ASIC for Deep Learning inference.



https://wccftech.com/nvidia-intros-drive-thor-balancing-ai-performance-for-full-vehicle-autonomy/

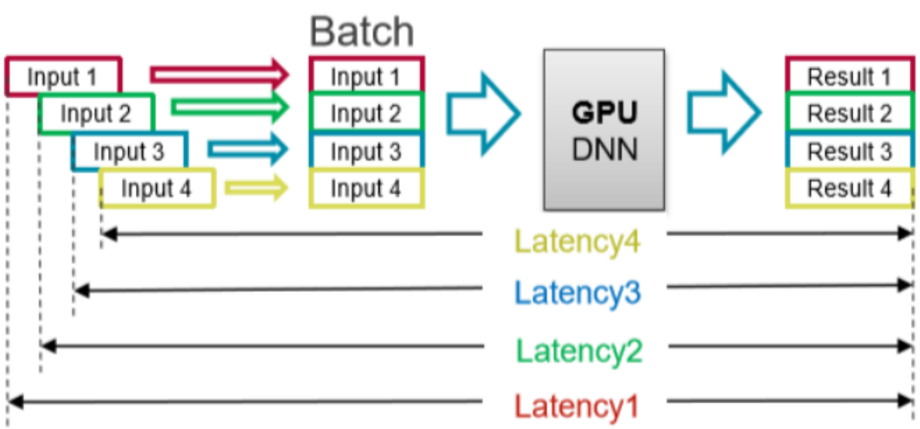
FPGAs

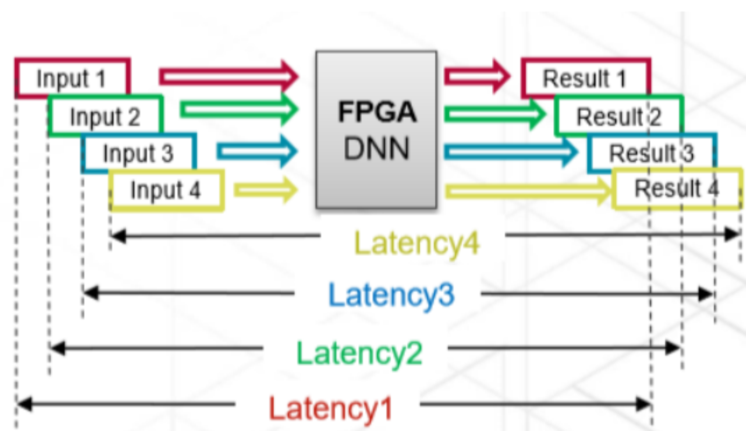
- FPGA is reprogrammable hardware, consisting of an array of Configurable Logic Blocks (CLBs) and interconnections which can be configured at design time or runtime.



FPGA vs. GPU

- FPGA has advantages over GPU for Deep Learning inference tasks.
- GPU performs computation in batches for efficient exploitation of SIMD (Single Instruction, Multiple Data) computation model.
  - This is ideally suited for training tasks, with well-known algorithms such as Stochastic Gradient Descent with mini-batches.
  - But not ideal for inference tasks.
    - Larger batch size leads to high throughput, but also high and nondeterministic latency for each data item.
    - Smaller batch size leads to low computation efficiency.
- FPGA can perform “batch-less” inference
  - Low and deterministic latency for any batch size.

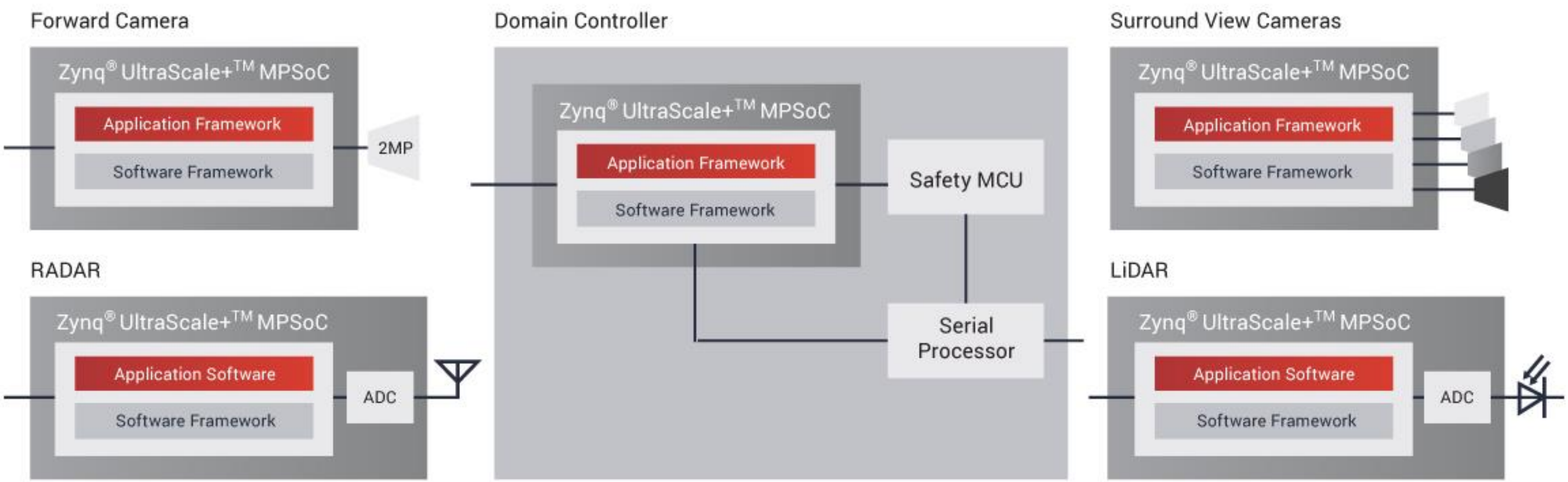




https://www.eetimes.com/where-do-fpgas-stand-in-auto-ic-race/#

FPGA for Automotive

- FPGAs can be integrated into sensors (camera, Lidar, radar), or serve as central compute engine in domain controller or AD computer
- Xilinx FPGAs have 90% market share in lidars.
- Intel is pushing hard into the automotive market, with acquisition of Altera in 2015.



https://www.xilinx.com/applications/automotive/automated-driving.html

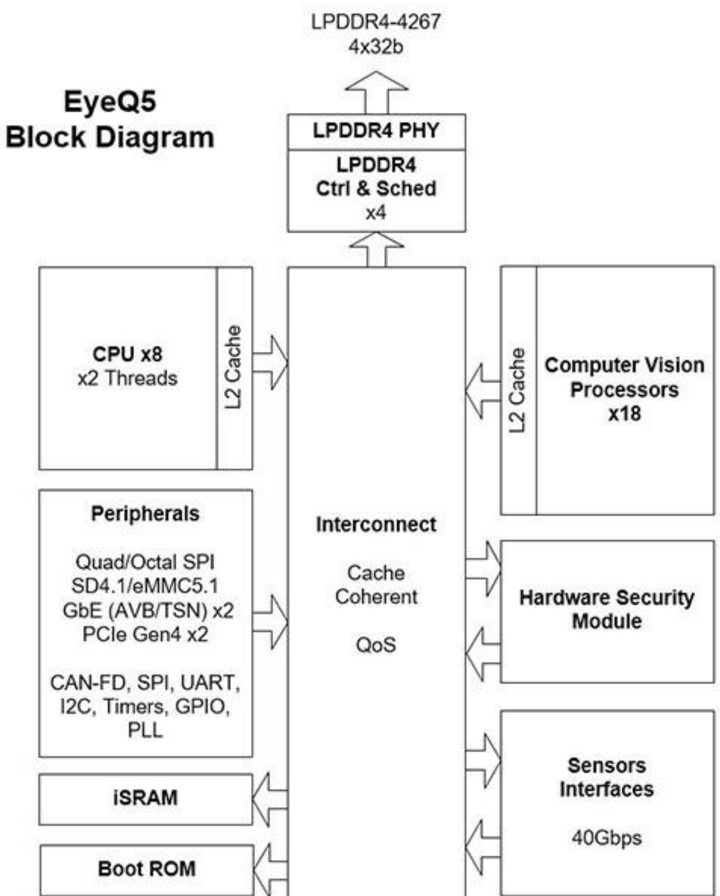
ASICs

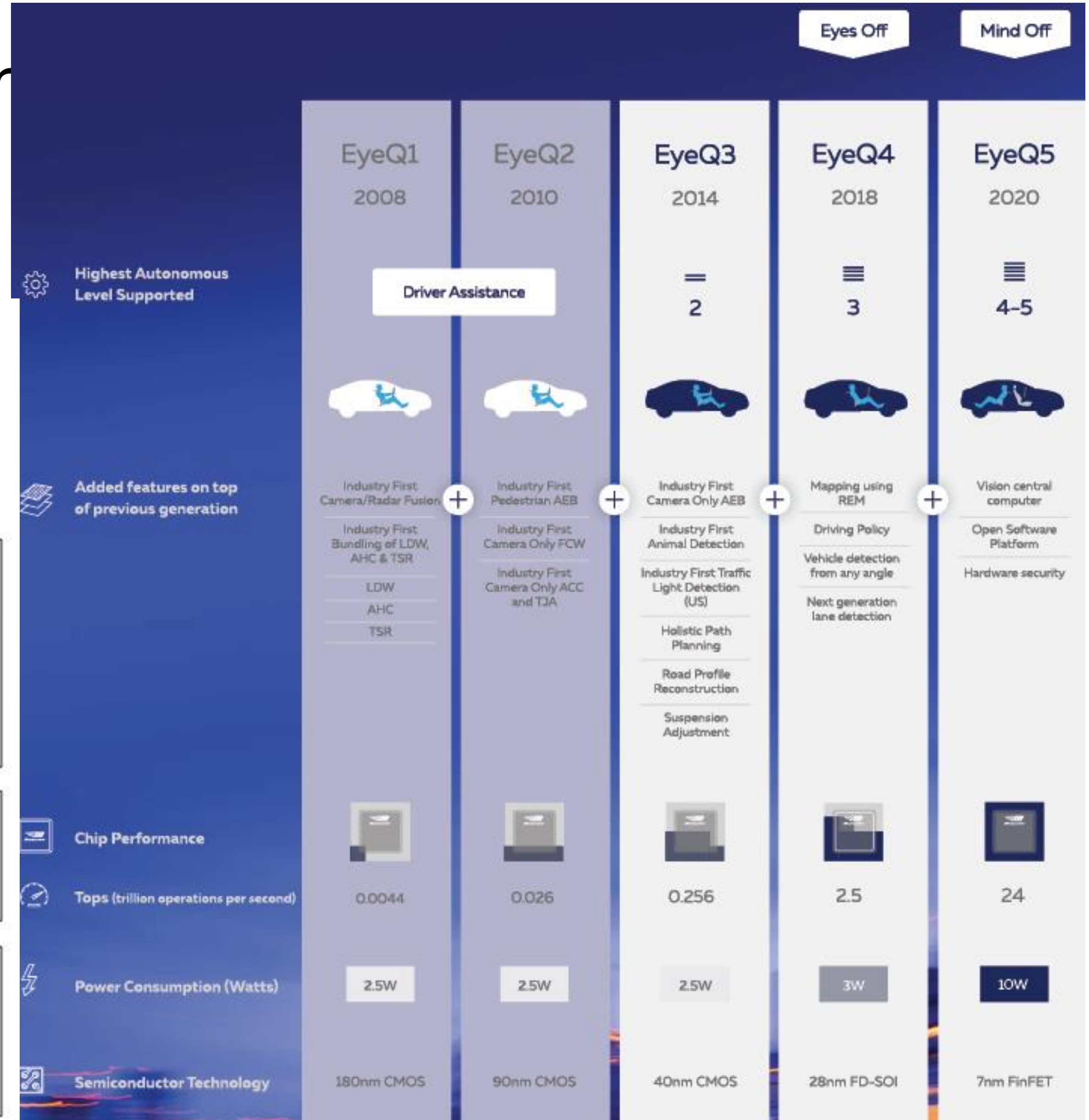
- ASICs for Deep Learning are often called Neural Processing Units or AI accelerators.
- ASICs, thanks to dedicated circuit design, may achieve up to 10x in computation efficiency and power consumption compared to CPU/GPU, and less dramatic, but still significant improvement compared to FPGA. The drawback is loss of programmability and flexibility.
  - Industry: almost every chip vendor provides some kind of AI accelerator, e.g. Google’s Tensor Processing Unit (TPU)
  - Academia: AI accelerators is a dominating topic in top conferences in computer architecture, including ISCA, MICRO and HPCA.

MobileEye

EyeQ Series

- “Computer Vision Processors” are ASICs

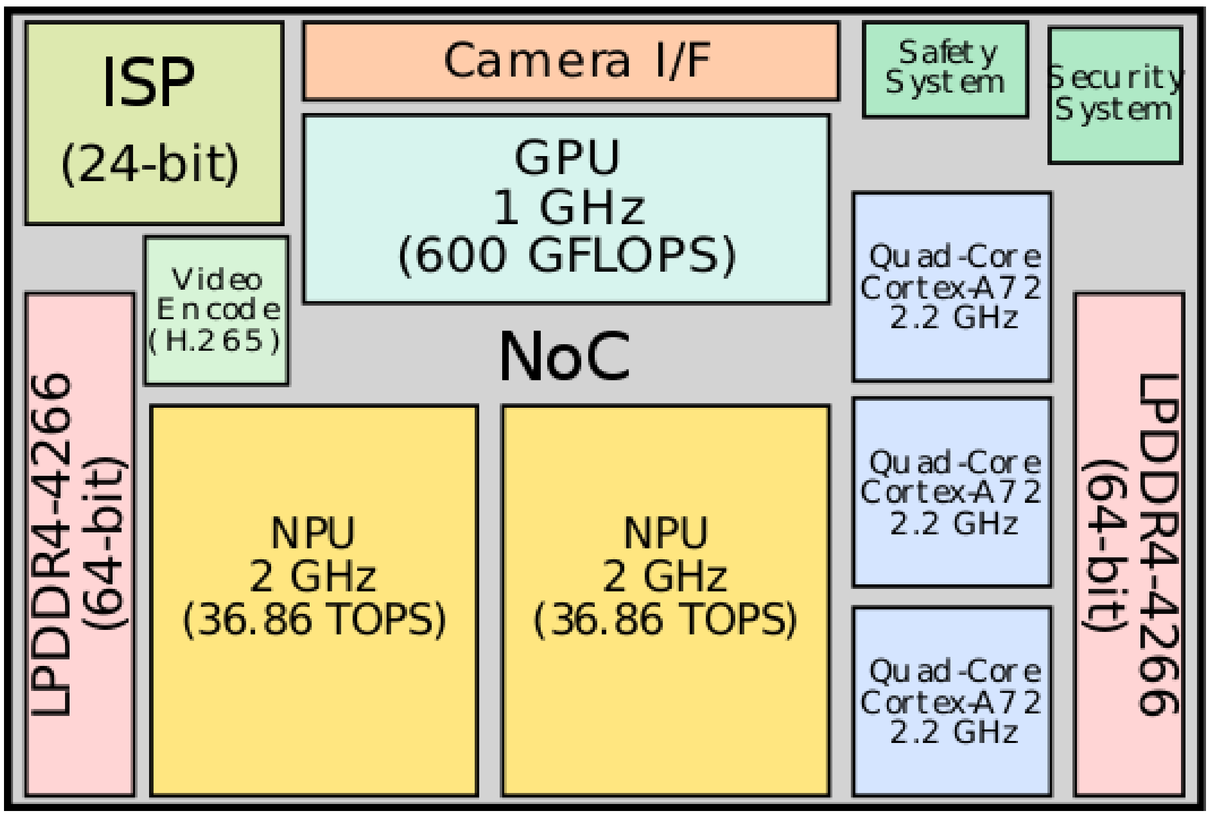




https://www.mobileye.com/our-technology/evolution-eyeq-chip/

Tesla FSD

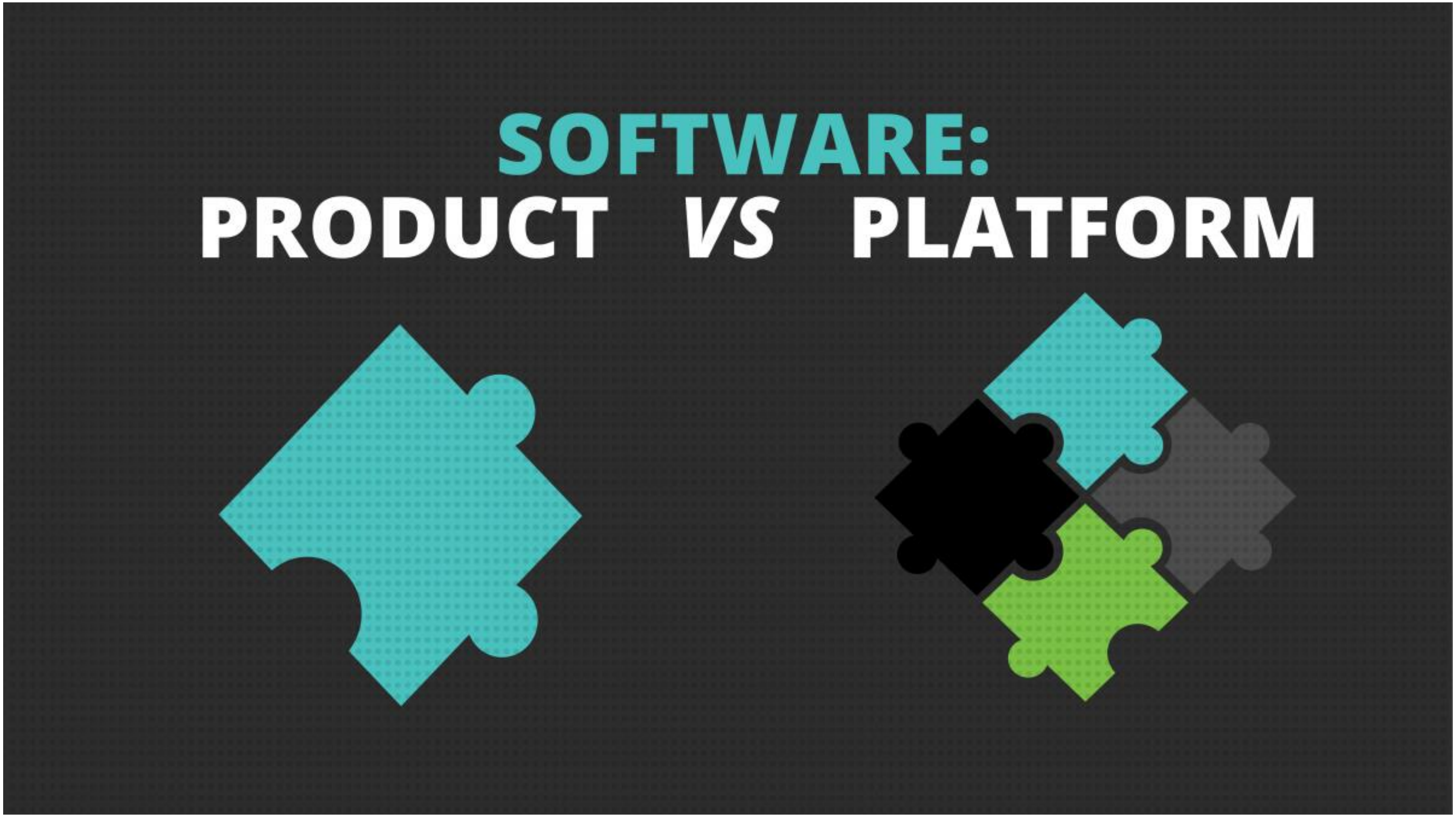
- Full Self-Driving Chip (FSD) is designed by Tesla and introduced in early 2019 for their own cars.
- It incorporates 3 quad-core Cortex-A72 clusters for a total of 12 CPUs operating at 2.2 GHz, a GPU operating 1 GHz, 2 Neural Processing Units (NPUs) operating at 2 GHz, and various other hardware accelerators.
  - NPU: ASIC for Deep Learning inference



Products from Automotive OEMs and Suppliers

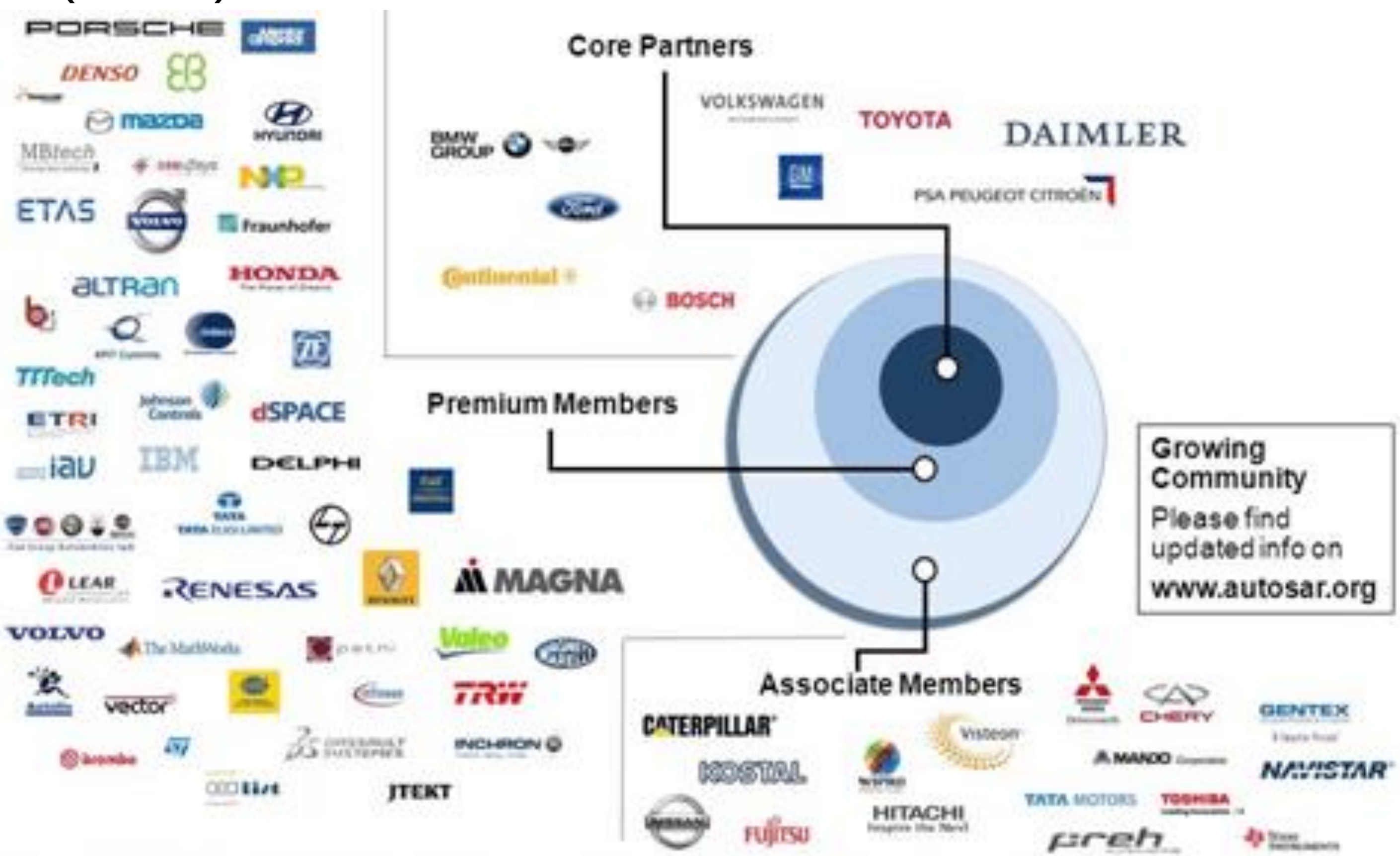
- These companies do not design chips. Rather, they offer integration solutions based on products from chip vendors like NVIDIA.
  - Delphi/Audi zFAS (zentrales Fahrerassistenz-Steuergeraet)
    - based on NVIDIA Tegra K1 and Mobileye EyeQ3
    - Hedging their bet on products from two mortal enemies ☺
  - ZF ProAI
    - based on NVIDIA DRIVE PX2
  - Bosch AI Car Computer
    - based on NVIDIA DRIVE AGX Xavier
  - Continental ADCU; Visteon DriveCore; NXP BlueBox; Renesas R-Car…

SOFTWARE PLATFORMS

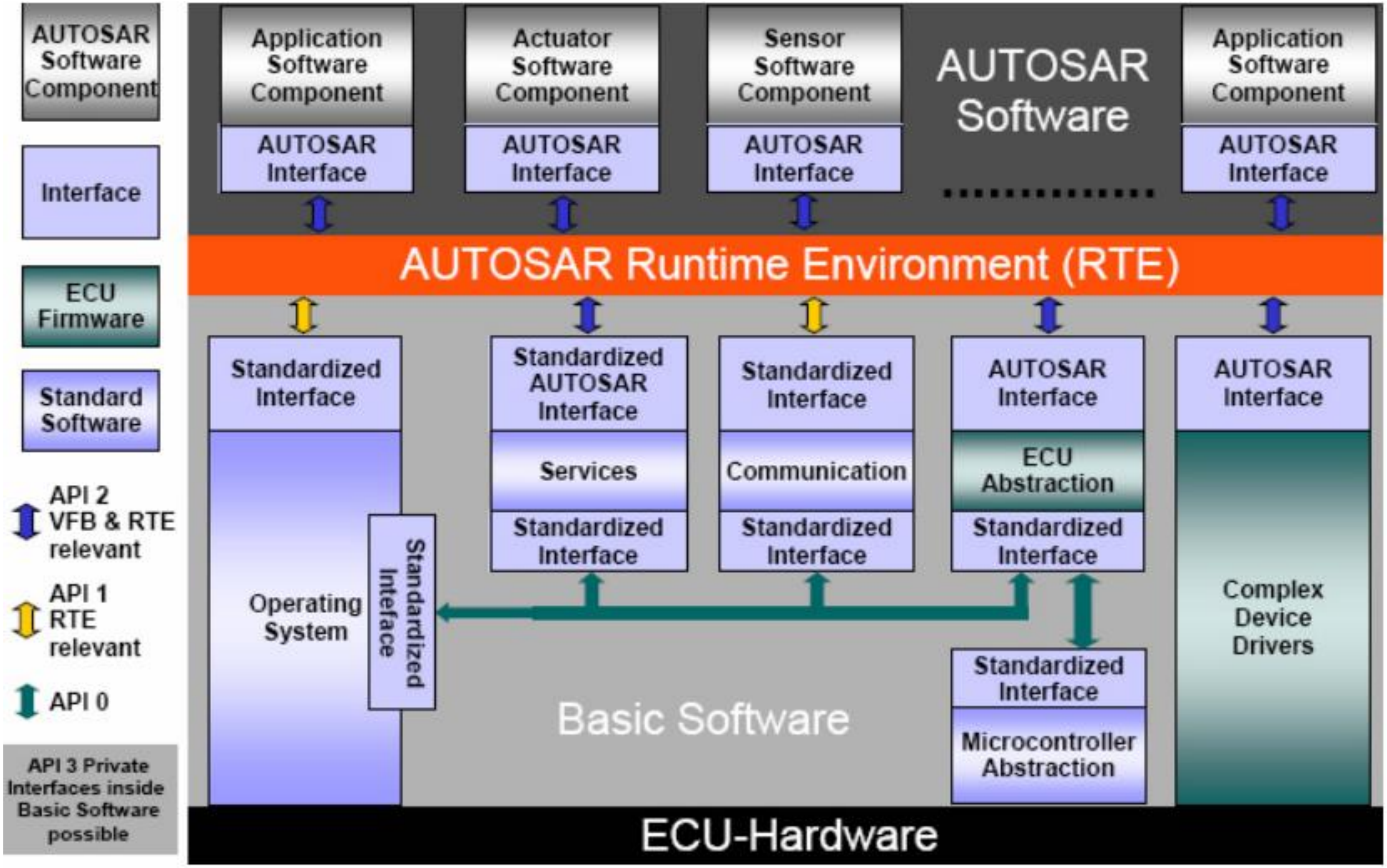


The AUTOSAR Consortium

- AUTomotive Open System ARchitecture (AUTOSAR) is a global development partnership of automotive interested parties founded in 2003. It pursues the objective to create and establish an open and standardized software architecture for automotive Electronic Control Units (ECUs).

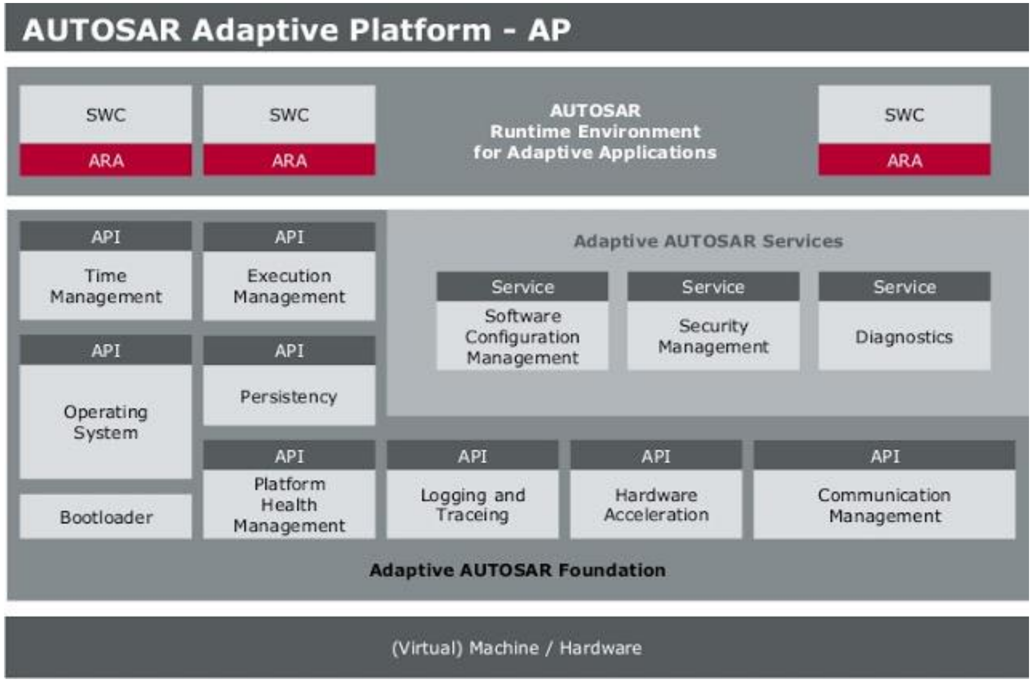


AUTOSAR Classic Platform



AUTOSAR Adaptive Platform

- AUTOSAR-AP is an industry standard that specifies standard interfaces required for developing future high-performance multicore automotive ECUs.
- Compared to Classic AUTOSAR for resource-constrained safety-critical ECUs, AUTOSAR-AP is designed for high-performance ECUs, and allows dynamic linking of services and clients during ECU runtime, which facilitates Over-the-Air (OTA) Update.



Integration of Multiple Software Platforms

- AUTOSAR CP (labeled C) is used for safety-critical ECUs for low-level control and interfacing with actuators
- AUTOSAR AP (labeled A) is used for high-performance AD computer.
- Non-AUTOSAR (labeled N) may be Linux or Android, for non-safety-critical IVI (In-Vehicle Infotainment) and COTS (Commercial Off-the-Shelf) applications.

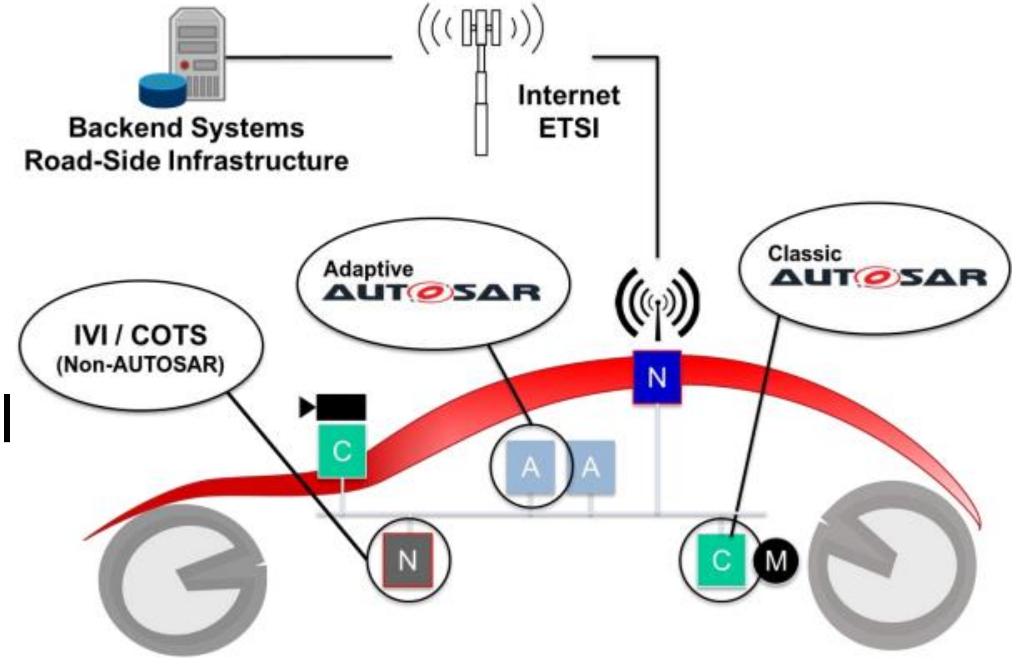


Figure 2-1 Exemplary deployment of different platforms

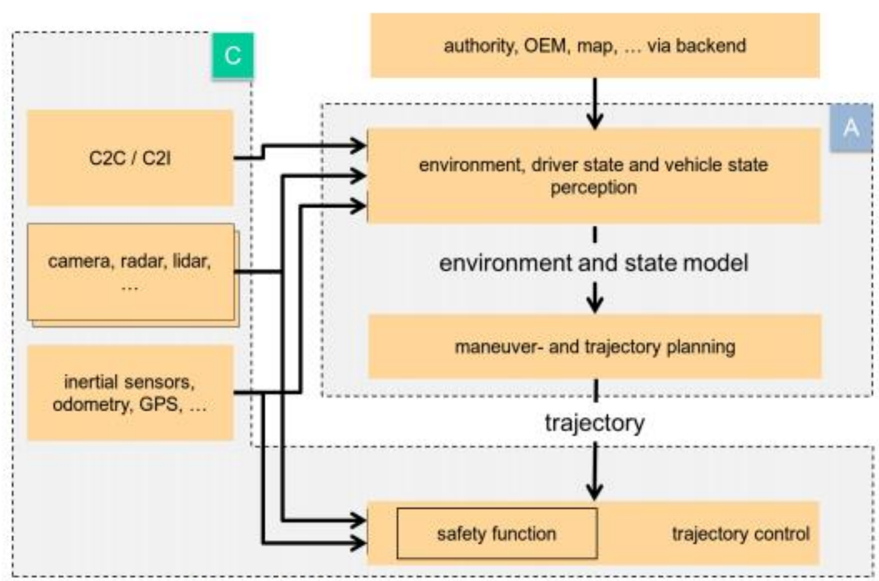
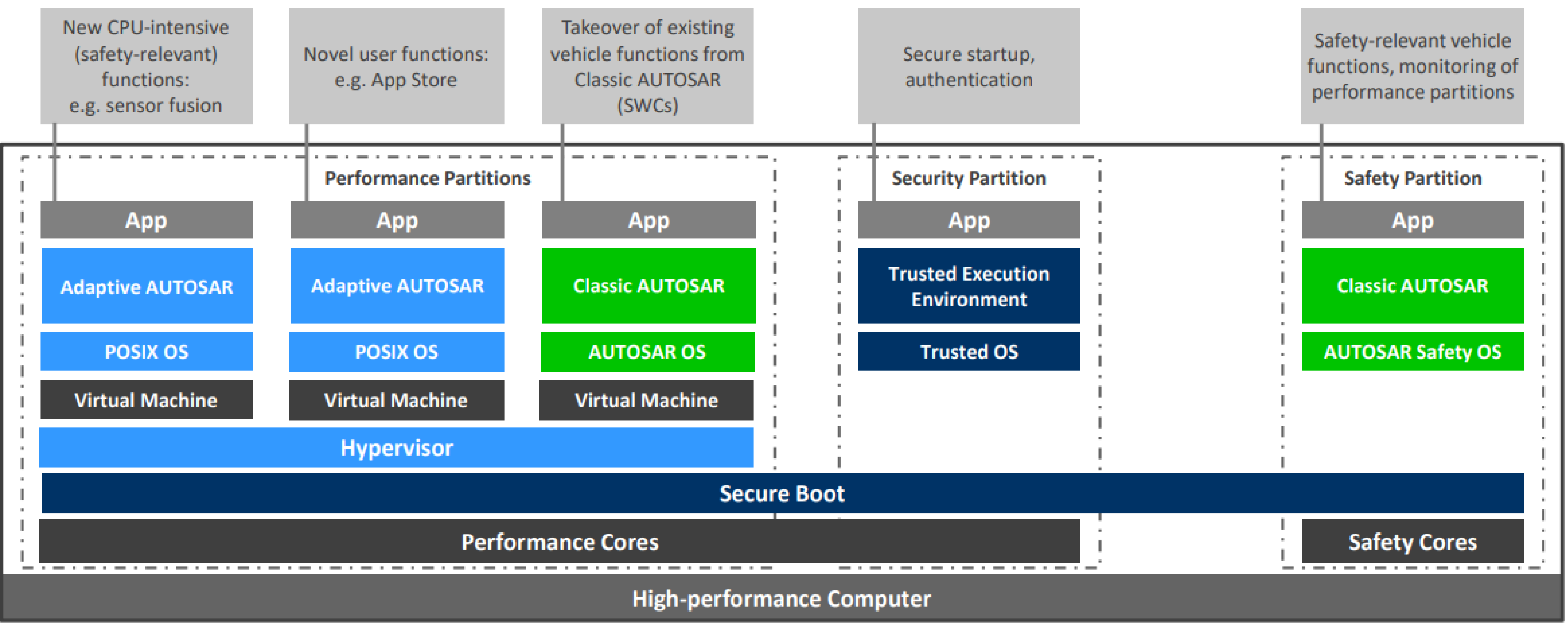


Figure 2-2 Exemplary interactions of AP and CP

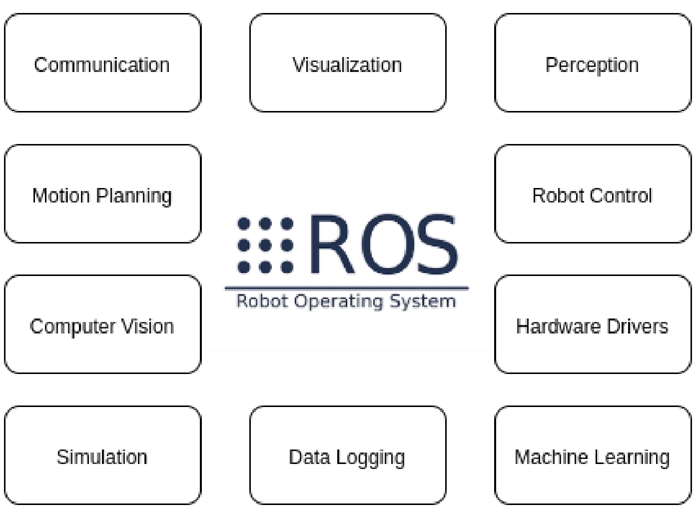
EB corbos

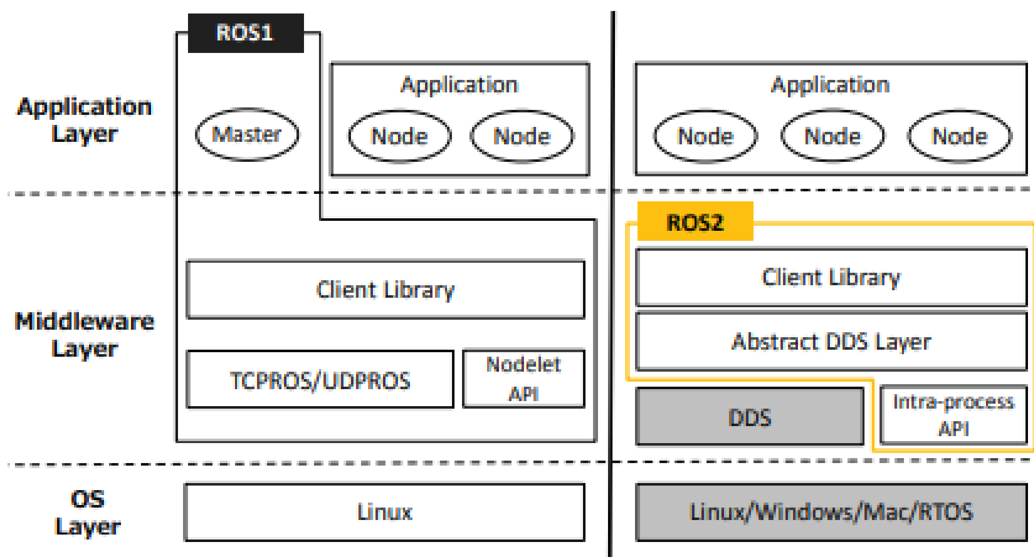
- EB corbos is a software platform that uses virtualization technology (hypervisor) to integrate AUTOSAR AP, CP and non-AUTOSAR OS on a single multicore ECU, while achieving high degree of isolation between different Virtual Machine partitions, including:
  - Performance Partitions with complex Performance Cores for high-performance, subject to low-levels of safety certification
  - Safety Partition with simpler Safety Cores for safety-critical functions, subject to high-levels of safety certification
  - Security Partition with processor security extensions (e.g., ARM TrustZone, Intel Software Guard Extensions (SGX)) for secure boot, crypto operations, etc.
- An example of a Mixed-Criticality System, where subsystems with different safety criticality levels are integrated on the same platform



Robot Operating System (ROS)

- ROS is a set of software libraries and tools for building robotic applications. Many companies use ROS to develop AVs. It uses the publish-subscribe paradigm for inter-node communication.
  - ROS has a Master node that provides naming and registration services to the rest of the nodes.
  - ROS 2 removed the Master node, and uses publish-subscribe middleware DDS (Data Distribution Service).
- A drawback of ROS compared to AUTOSAR:
  - Since ROS uses Linux as the OS, it is not possible to pass high-level of safety certification (ASIL-D).

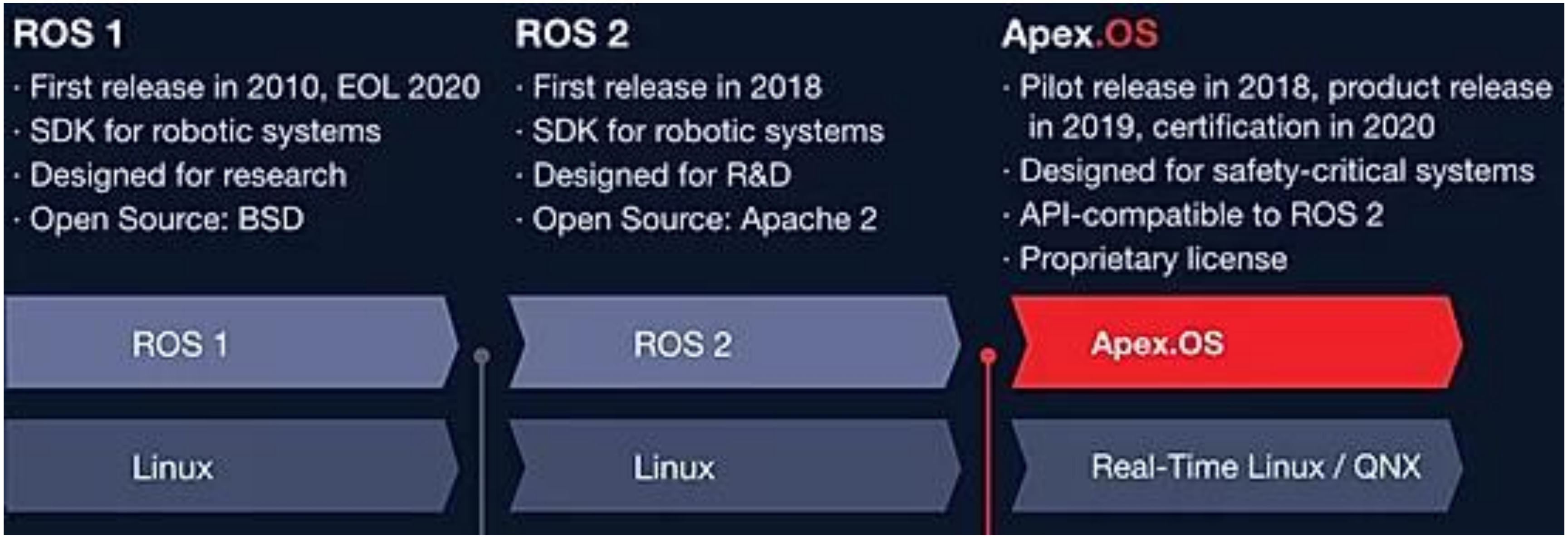




Maruyama Y, Kato S, Azumi T. Exploring the performance of ROS2[C]//Proceedings of the 13th International Conference on Embedded Software. 2016: 1-10.

Apex.OS

- "Safe and certified software framework for autonomous mobility systems."
- Aims to be certified as a Safety Element out of Context (SEooC) up to ASIL D.
  - Hard real-time, static memory allocation (no new() or malloc()), callbacks vs. waitset, security, testing, real-time I/O logging…
  - Real-Time Linux or QNX as the RTOS.



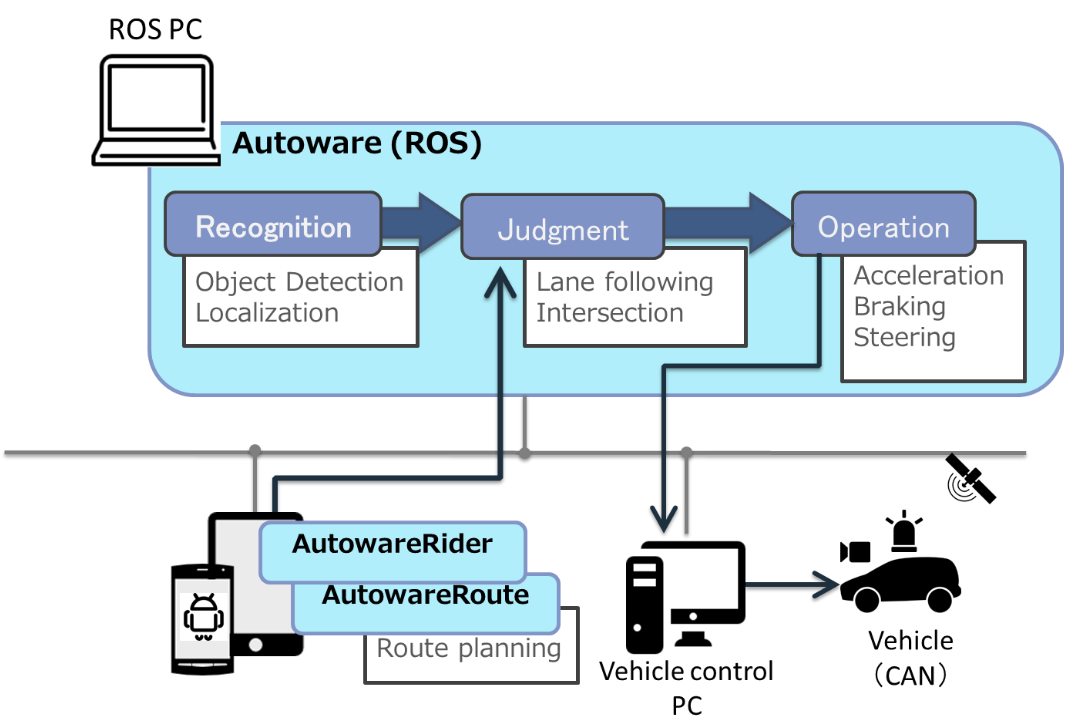
NVIDIA DRIVE Software Framework

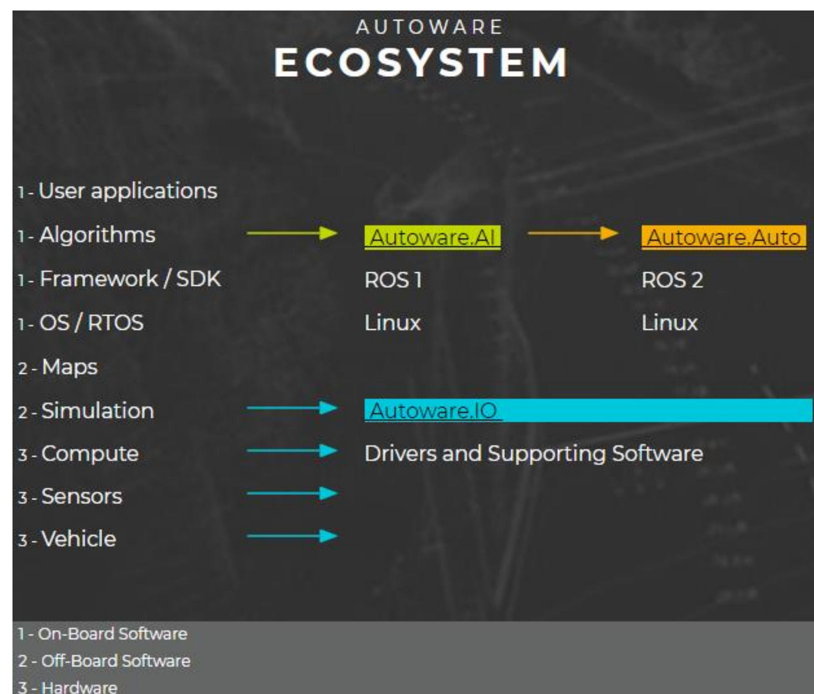
NVIDIA DRIVE Software Framework

- An open-source framework for AD (only for NVIDIA hardware).
  - **DRIVE OS** is a foundational software stack consisting of an embedded Real Time OS (RTOS), hypervisor, CUDA libraries, Tensor RT, and other modules that give you access to the hardware engines.
  - **DriveWorks SDK** enables developers to implement AV solutions by providing a comprehensive library of modules, developer tools, and reference applications.
  - **DRIVE AV** provides perception, mapping, and planning modules that utilize the DriveWorks SDK.
  - **DRIVE IX** provides full cabin interior sensing capabilities needed to enable AI cockpit solution.

Autoware

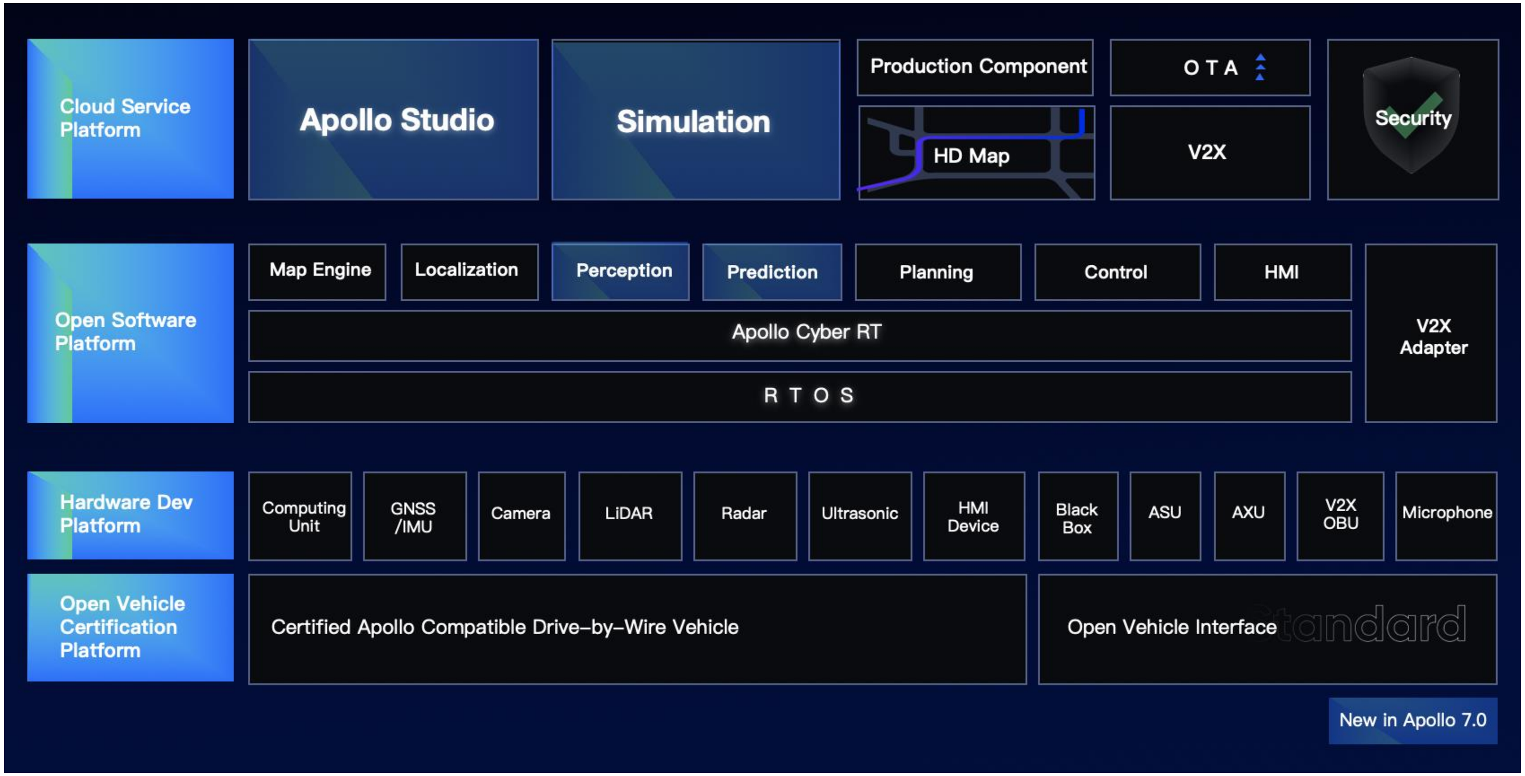
- Open-source AD platform from Japan.
  - Autoware.AI (https://www.autoware.ai) is based on ROS-1.
  - Autoware.auto (https://www.autoware.auto) is the new version based on ROS2.



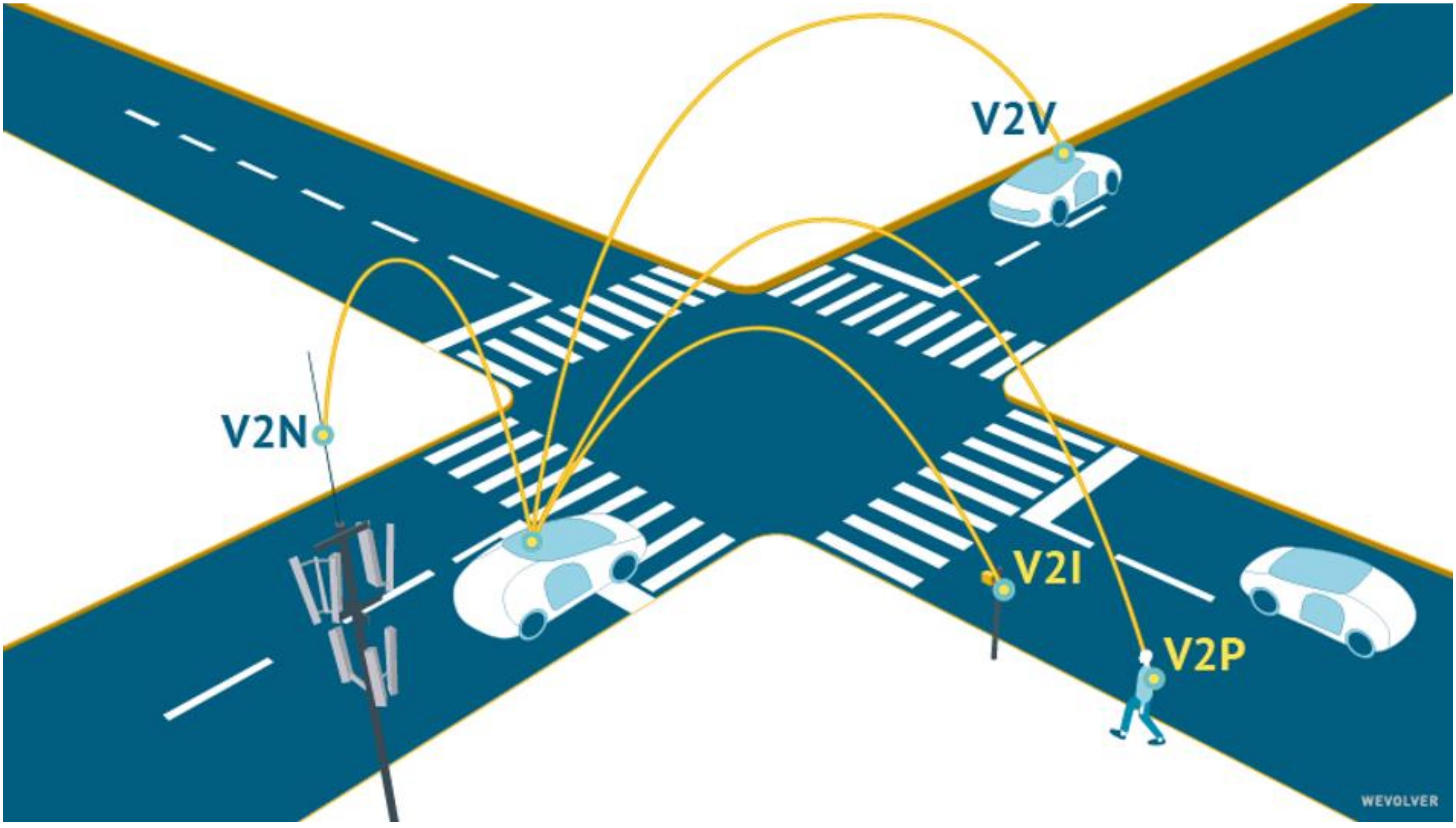


Baidu Apollo

- An open-source, hardware-neutral AD platform from China.
- Initially based on ROS, but later replaced ROS with their own components.
  - Real-Time Operating System (RTOS); Linux kernel with real-time patch
  - Cyber RT: lightweight, high-performance communication middleware

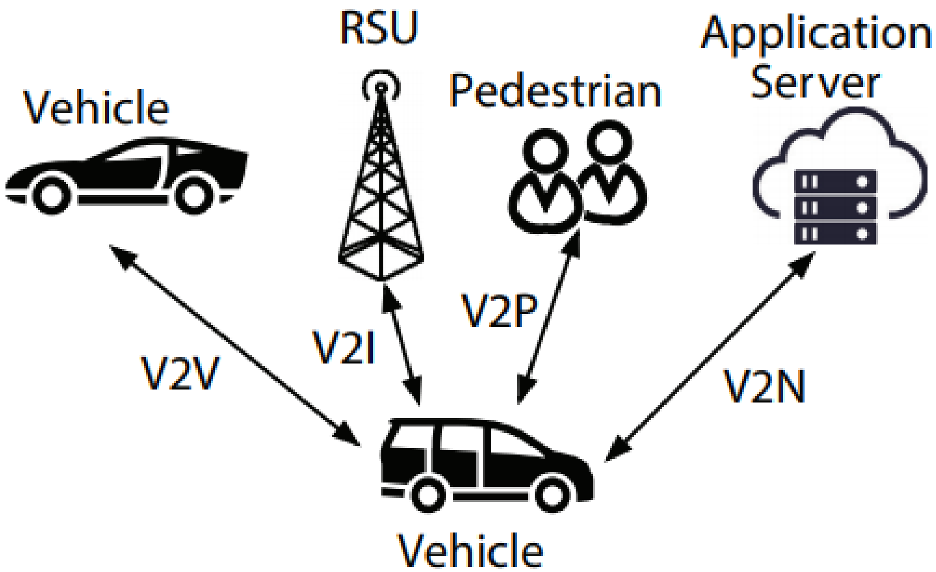


V2X



V2X Types

- Four types of V2X applications in 3GPP Standard:
  - Vehicle-to-Vehicle (V2V)
    - e.g., collision avoidance system
  - Vehicle-to-Pedestrian (V2P)
    - e.g., safety alerts to pedestrians and bicyclists
  - Vehicle-to-Infrastructure (V2I)
    - e.g., adaptive traffic light control, traffic-light optimal speed advisory
  - Vehicle-to-Network (V2N)
    - e.g., real-time traffic routing, cloud services
    - Also called Vehicle-to-Cloud



Two Camps

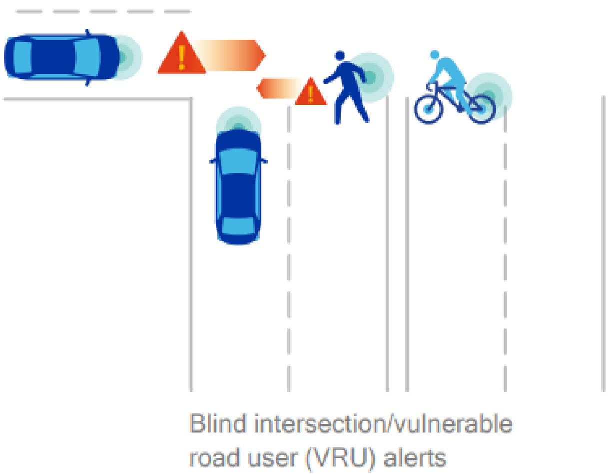
- DSRC (Distributed Short-Range Communication): Toyota, GM…
  - Based on WiFi, i.e., 802.11p at 5.9GHz
  - For latency-sensitive applications.
- C-V2X (Cellular V2X): Qualcomm/Ford…
  - Traditionally for latency-tolerant applications e.g., Over-the-Air (OTA) updates.
  - With the advent of 5G, C-V2X will become more prevalent, used also for latency-sensitive applications.
- C-V2X seems to be winning over DSRC in recent years. The following slides are based on Qualcomm's C-V2X approach

https://www.qualcomm.com/media/documents/files/accelerating-c-v2x-commercialization.pdf

V2X Applications in AD

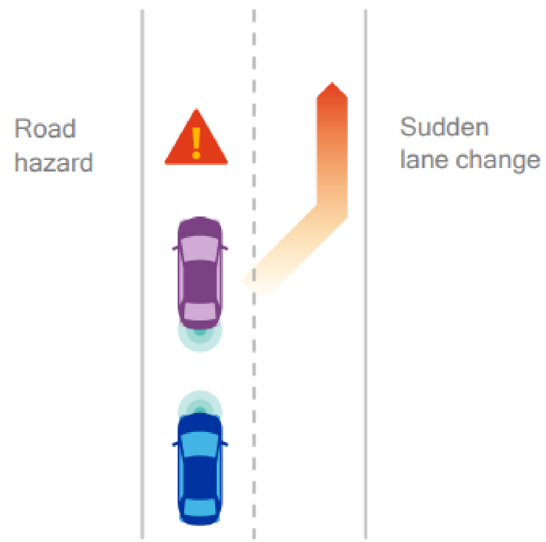
Non line-of-sight sensing

Provides 360˚ NLOS awareness, works at night and in bad weather conditions



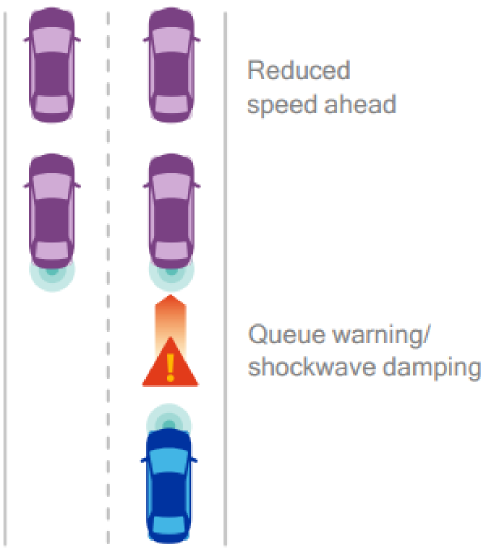
Conveying intent

Shares intent, sensor data , and path planning info for higher level of predictability



Situational awareness

Offers increased electronic horizon to support soft safety alerts and graduated warning



C-V2X Evolution towards 5G

- In 2020/07, 3GPP (3rd Generation Partnership Project) declared R16 to be frozen

	D2D communications	**Enhanced safety**	**Autonomous driving**
	R12/13	C-V2X R14 (Ph. I) C-V2X R15 (Ph. II)	C-V2X R16 5G NR support (Ph. III) (Advanced safety applications)
	Established foundation for basic D2D comm.	Enhanced communication's range and reliability for V2X safety	Ultra-reliable, low latency, high throughput communication for autonomous driving
Network independent	No	Yes	Yes
Communications[1]	Broadcast only	Broadcast only	Broadcast + Unicast/Multicast
High speed support	No	Yes	Yes
High density support	No	Yes	Yes
Throughput		High throughput for enhanced safety	Ultra-high throughput
Latency		Low latency for enhanced safety applications	Ultra-low latency
Reliability		Reliability for enhanced safety application	Ultra-high reliability
Positioning	No	Share positioning information	Wideband ranging and positioning

C-V2X Defines 2 Transmission Modes

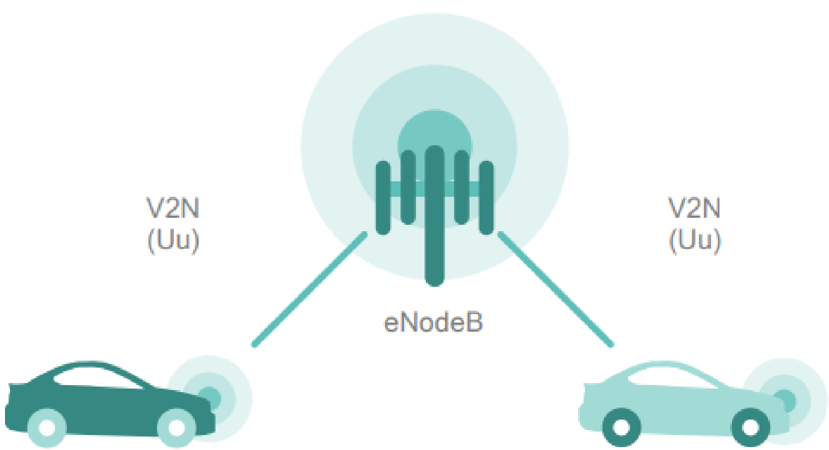
- V2X-Cellular: network communications going through base station (eNodeB)
- V2X-Direct: Device-to-Device direct communications without going through base station

Network communications

V2N on “Uu” interface operates in traditional mobile broadband licensed spectrum

Uu interface

e.g. accident 2 kilometer ahead

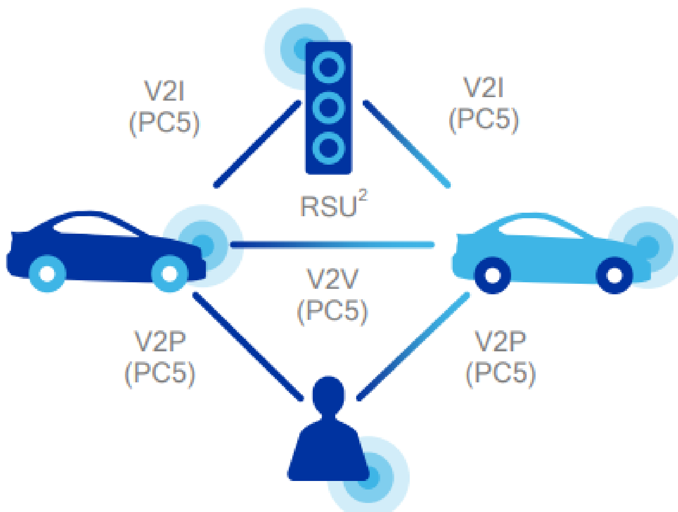


Direct communications

V2V, V2I, and V2P on “PC5” interface[1], operating in ITS bands (e.g. ITS 5.9 GHz) independent of cellular network

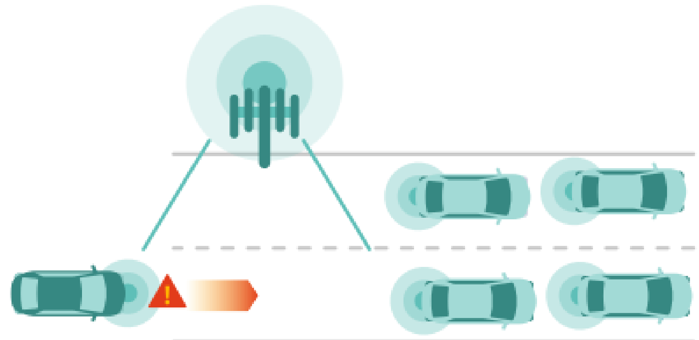
PC5 interface

e.g. location, speed

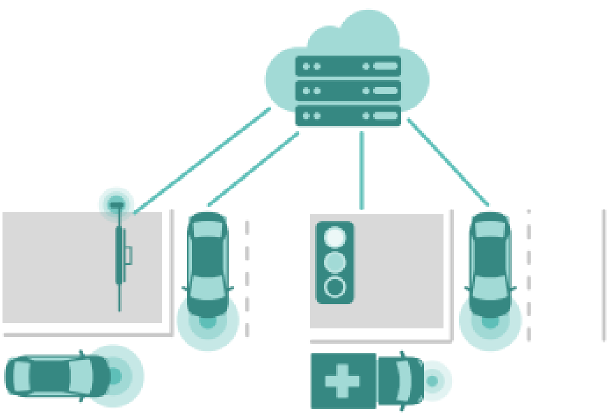


V2X-Cellular for Latency-Tolerant Use Cases

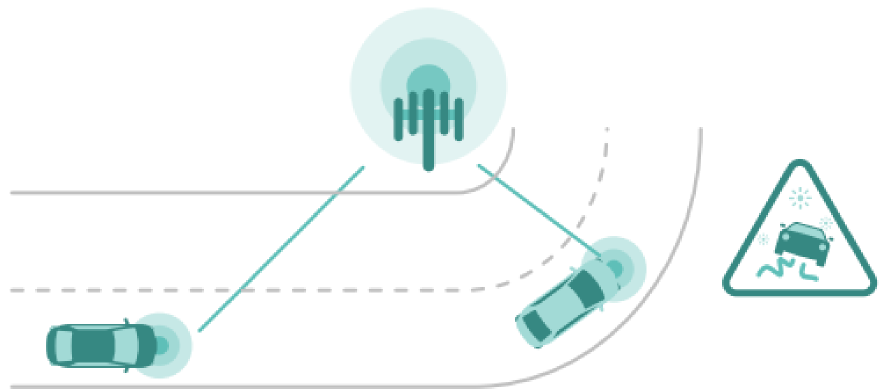
Discover parking and charging



Traffic flow control/ Queue warning



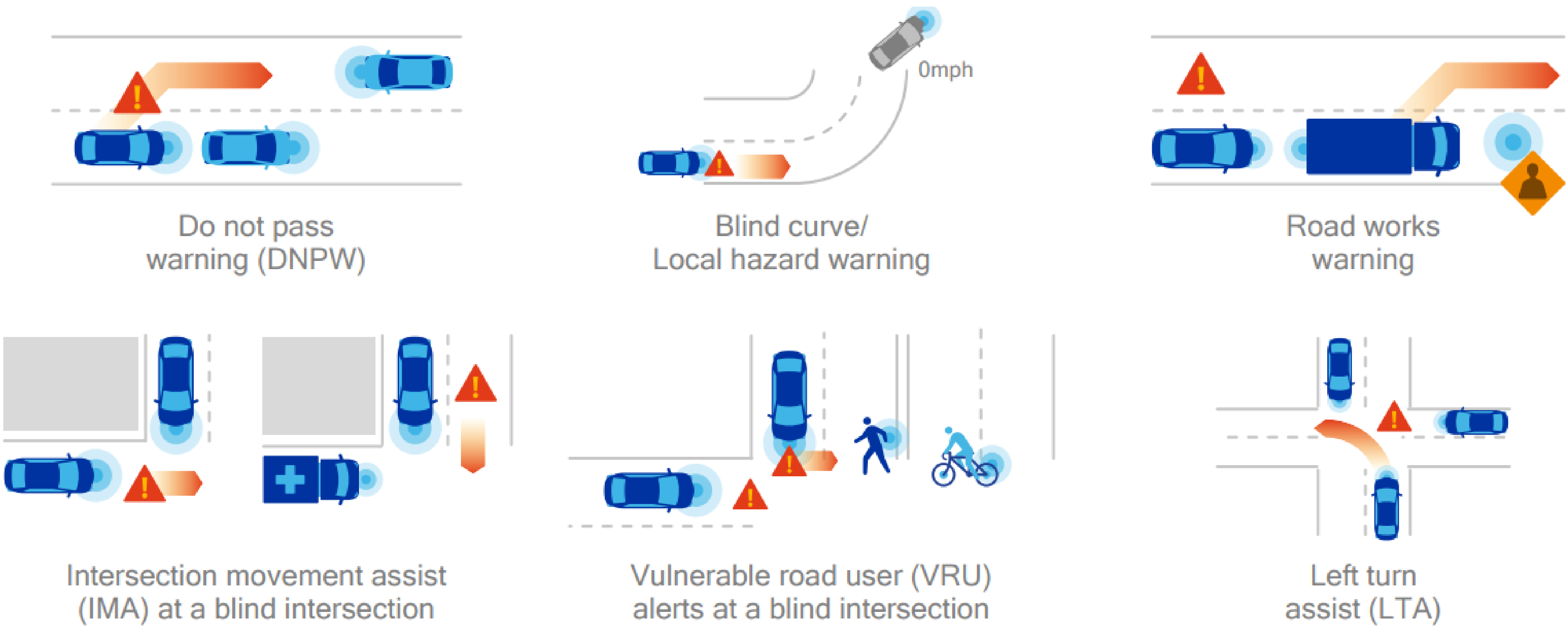
Cloud-based sensor sharing



Road hazard warning 1 km ahead

V2X-Direct for Latency-Sensitive Active Safety Use Cases

- Useful for NLOS (Non-Line-of-Sight) scenarios.

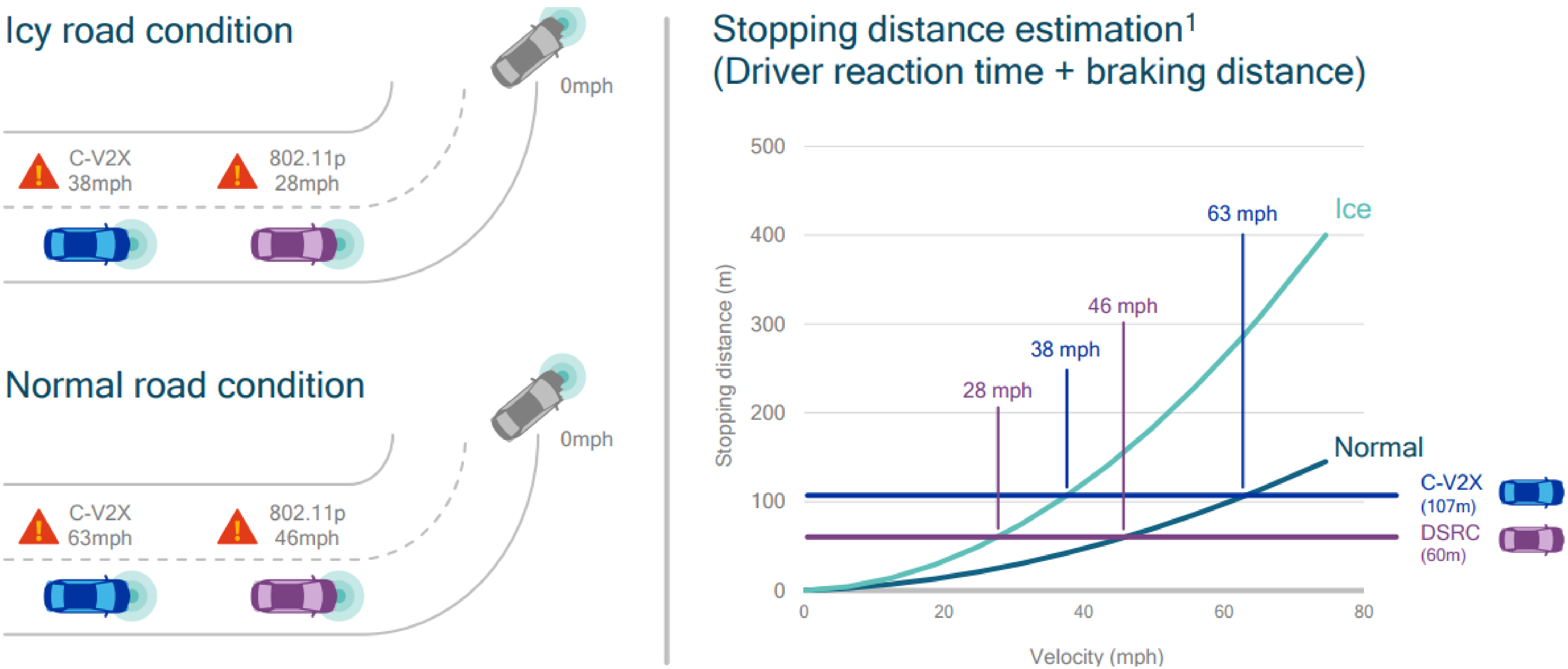


C-V2X Deployment

- Combined RSUs (Road-Side Units) with direct link (PC5) interface for V2X-Direct, and 4G/5G base stations for V2X-Cellular, benefiting from cellular network densification in 5G (smaller and denser cells)

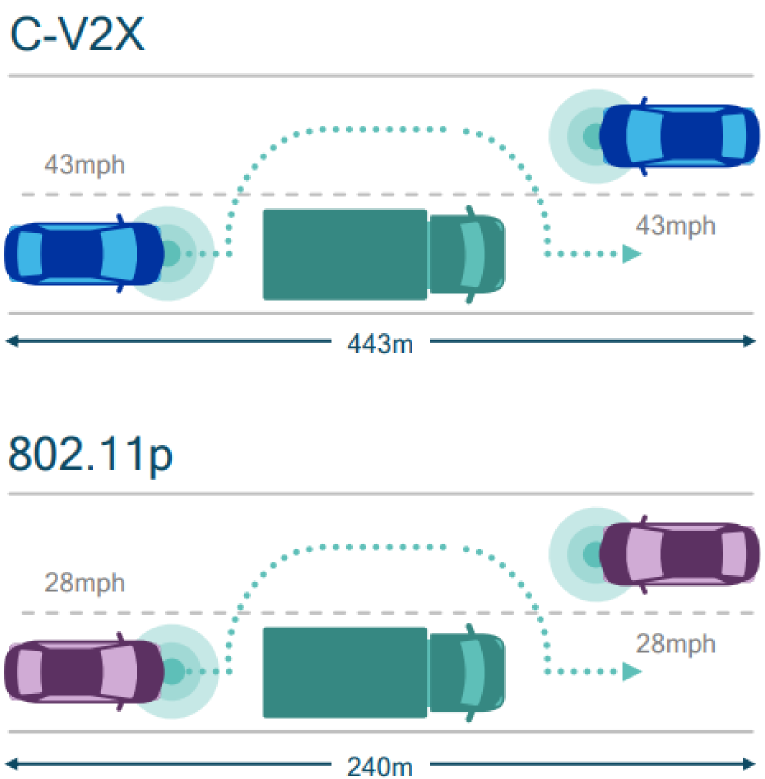
Use Case: Disabled Vehicle after Blind Curve

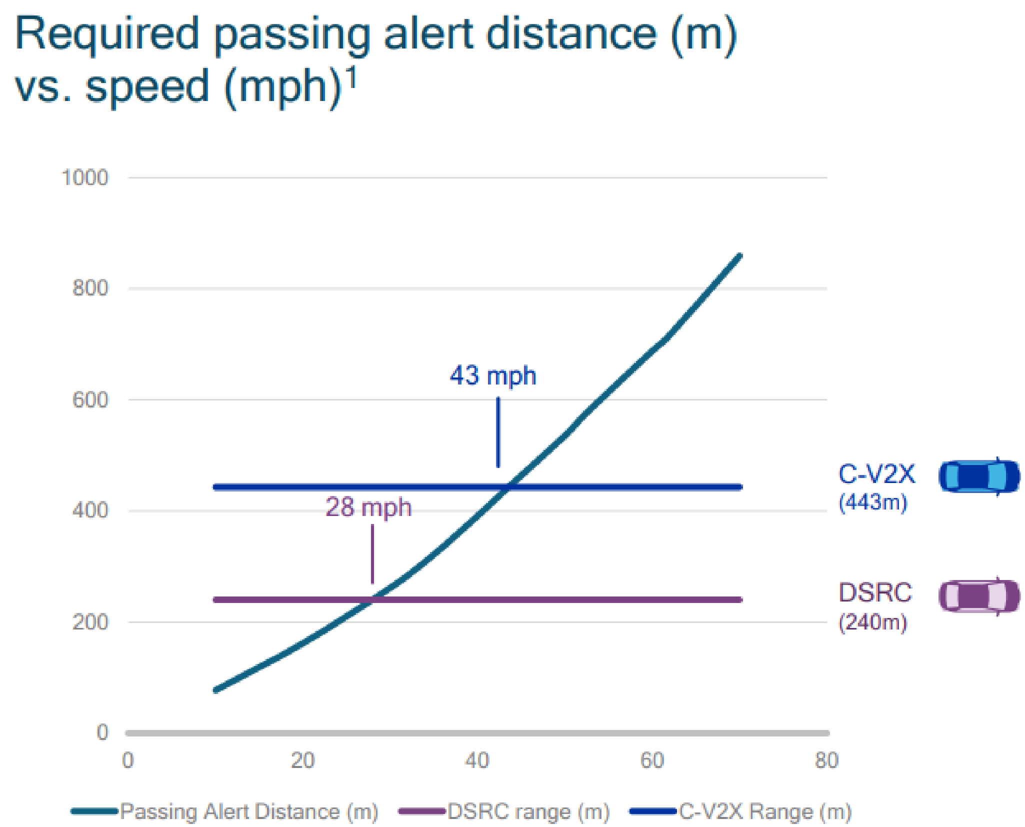
- C-V2X has (at least 2x) longer range than DSRC/802.11p, which enables the ego-vehicle to get warning message earlier, hence travel at higher speed while avoiding collision with the disabled vehicle



Use Case: Do Not Pass Warning

- C-V2X's longer range enables the ego-vehicle to get warning message earlier, hence travel at higher speed while avoiding collision with the disabled vehicle





Industry Consortium

- 5GAA is a cross-industry consortium that defines 5G V2X communications

5GAA Automotive Association

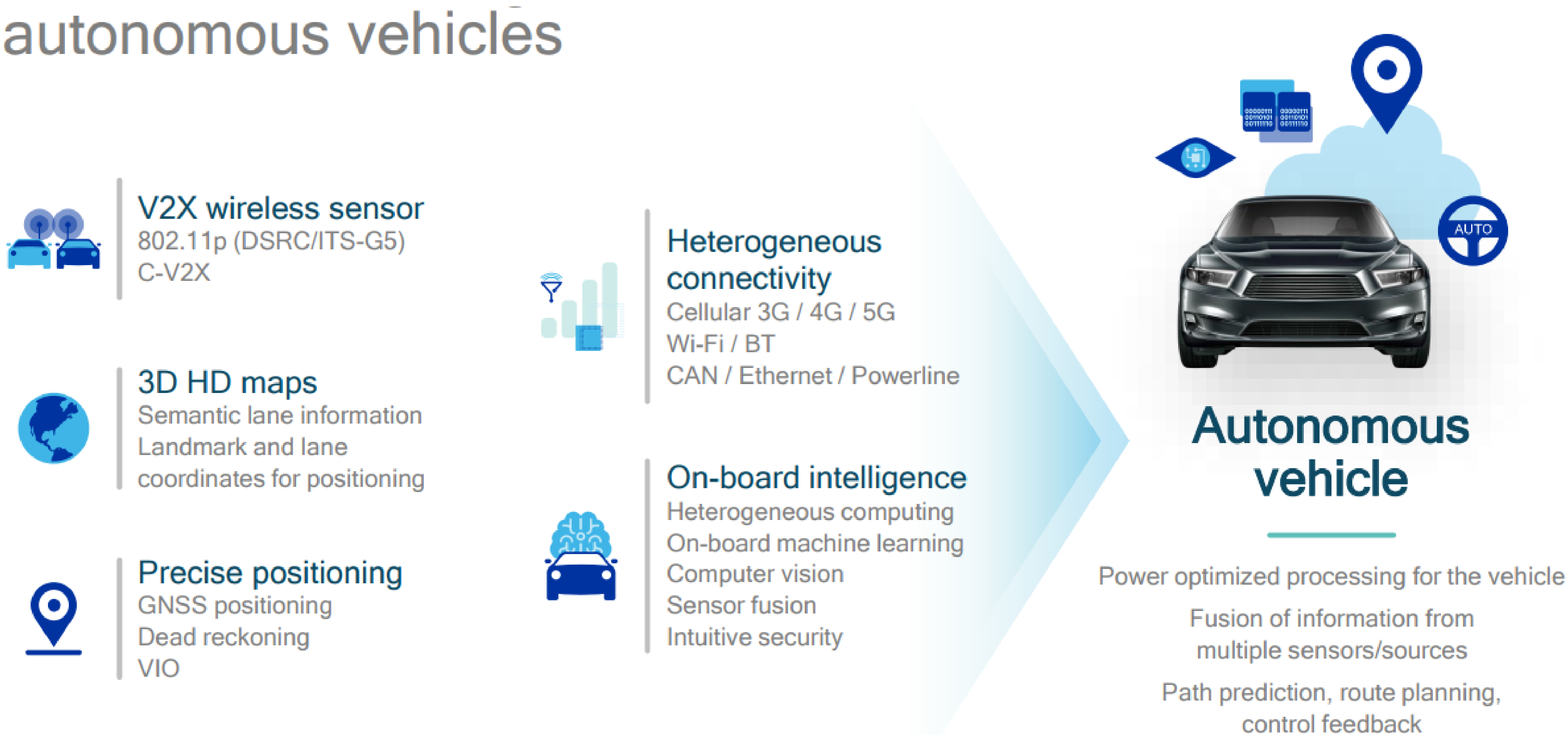
Automotive industry
Vehicle platform, hardware, and software solutions

Telecommunications
Connectivity and networking systems, devices, and technologies

End-to-end solutions for intelligent transportation mobility systems and smart cities

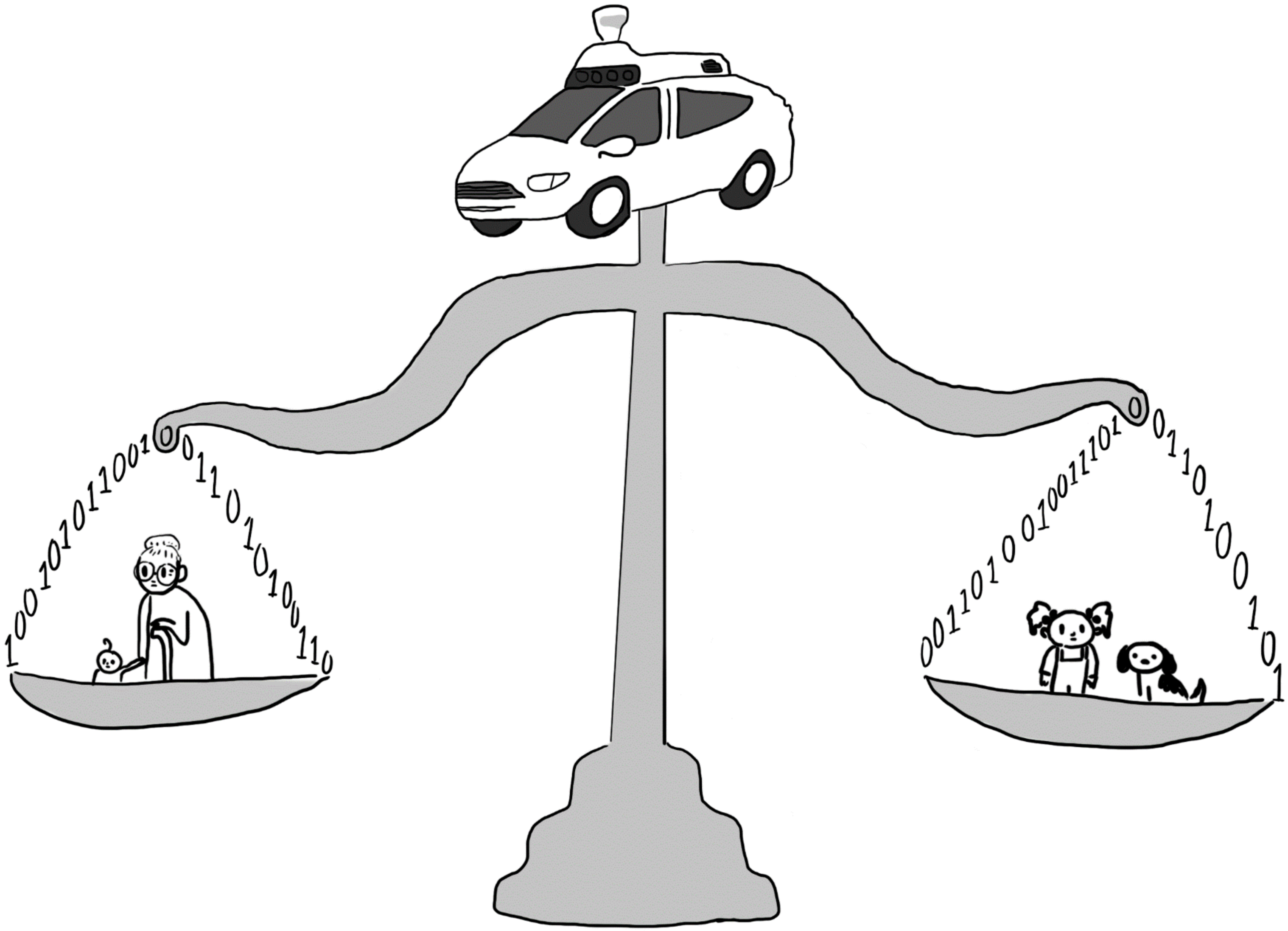
Analog Devices AT&T Audi BAIC BMW Bosch CAICT CETECOM China Mobile Continental Daimler
Danlaw Denso Ericsson FEV Ficosa Ford Gemalto Hirschmann Car Communication Huawei Infineon
Intel Interdigital Jaguar KDDI Keysight Technologies KT Laird Land Rover LG MINI muRata Nokia
NTT DoCoMo P3 Panasonic Qualcomm Rohde & Schwarz ROHM Rolls-Royce SAIC Motor Samsung Savari
SK Telecom SoftBank T-Mobile Telefonica Telstra TÜV Rheinland Valeo Verizon VLAVI Vodafone ZF ZTE

5G Accelerates AVs



- My thoughts: Massive deployment of V2X is necessary for AVs to benefit from V2X; It is a promising technology, but current AV players are not counting on V2X.

ETHICAL ISSUES

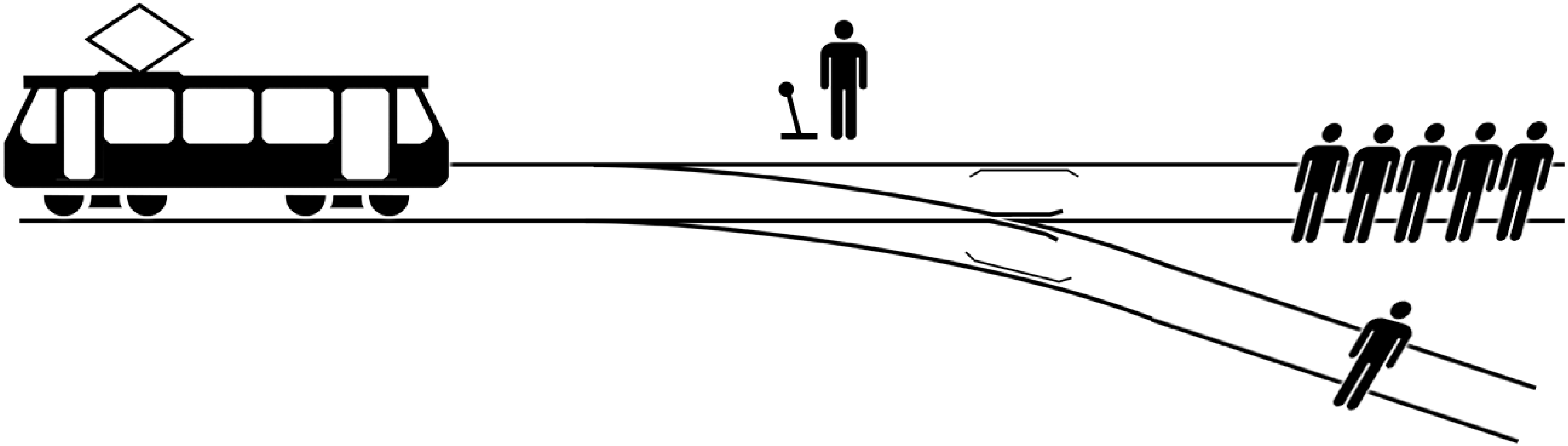


Two Introductory Videos

- The ethical dilemma of self-driving cars - Patrick Lin
  - https://www.youtube.com/watch?v=ixIoDYVfKA0
- Moral Machines: How culture changes values
  - https://www.youtube.com/watch?v=jPo6bby-Fcg

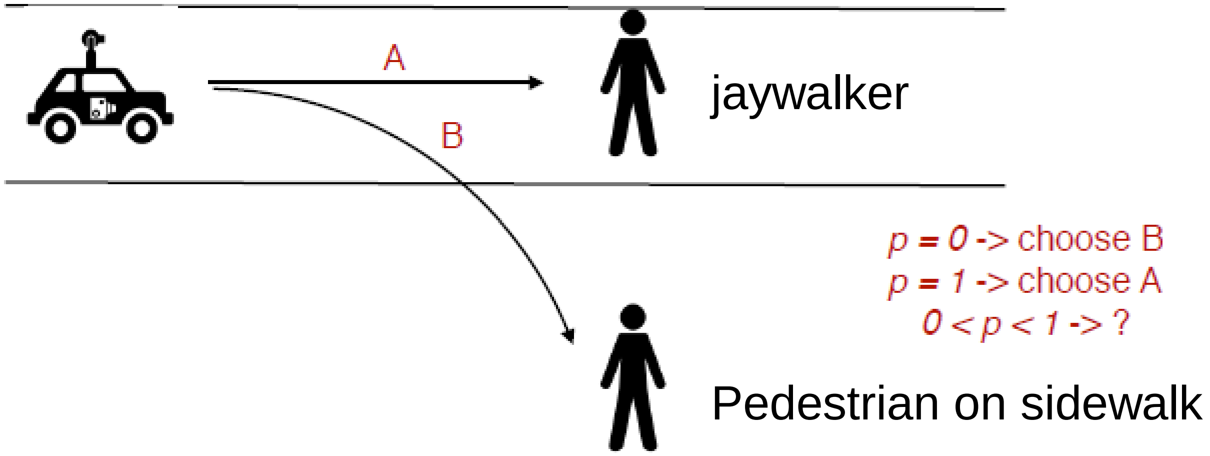
The Trolley Problem

- There is a runaway trolley barreling down the railway tracks. Ahead, on the tracks, there are five people tied up and unable to move. The trolley is headed straight for them. You are standing some distance off in the train yard, next to a lever. If you pull this lever, the trolley will switch to a different set of tracks. However, you notice that there is one person on the side track. You have two options:
  - Do nothing and allow the trolley to kill the five people on the main track.
  - Pull the lever, diverting the trolley onto the side track where it will kill one person.
- What is the right thing to do?



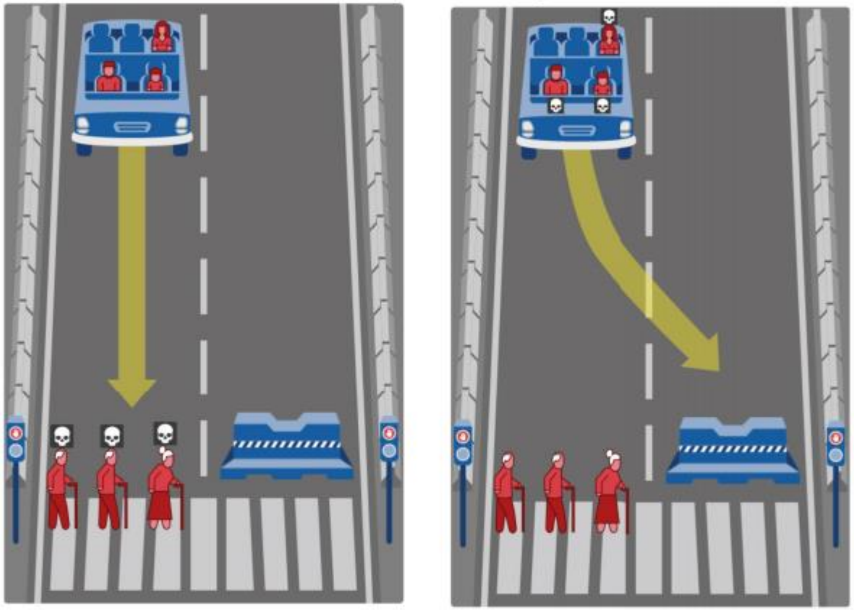
Variant of Trolley Problem with a Probability Threshold

- You are in a situation where:
  - A. you kill a pedestrian with probability 1, but it's not your fault
  - B. you kill a different pedestrian with probability p, and it is your fault
- What is your threshold value p_{th} for making the choice?
  - if($p \geq p_{th}$) choose A; otherwise choose B



MIT Moral Machine Experiment

- A 2016 survey indicates that people wanted an autonomous vehicle to protect pedestrians even if it meant sacrificing its passengers — but also that they wouldn't buy self-driving vehicles programmed to act this way. This prompted the MIT Moral Machine Experiment, a platform for gathering a human perspective on moral decisions made by machine intelligence, such as AVs (http://moralmachine.mit.edu/)
- An AV must choose between killing two passengers or five pedestrians. An AV experiences a sudden brake failure. Staying on course would result in the death of two elderly men and an elderly woman who are crossing on a 'do not cross' signal (left). Swerving would result in the death of three passengers: an adult man, an adult woman, and a boy (right).
- You can also design other scenarios. Accident scenarios are generated with nine factors: sparing humans (versus pets), staying on course (versus swerving), sparing passengers (versus pedestrians), sparing more lives (versus fewer lives), sparing men (versus women), sparing the young (versus the elderly), sparing pedestrians who cross legally (versus jaywalking), sparing the fit (versus the less fit), and sparing those with higher social status (versus lower social status).
- This platform gathered 40 million decisions in ten languages from millions of people in 233 countries



Awad, Edmond, et al. "The moral machine experiment." *Nature* 563.7729 (2018): 59-64.

Individual Variations

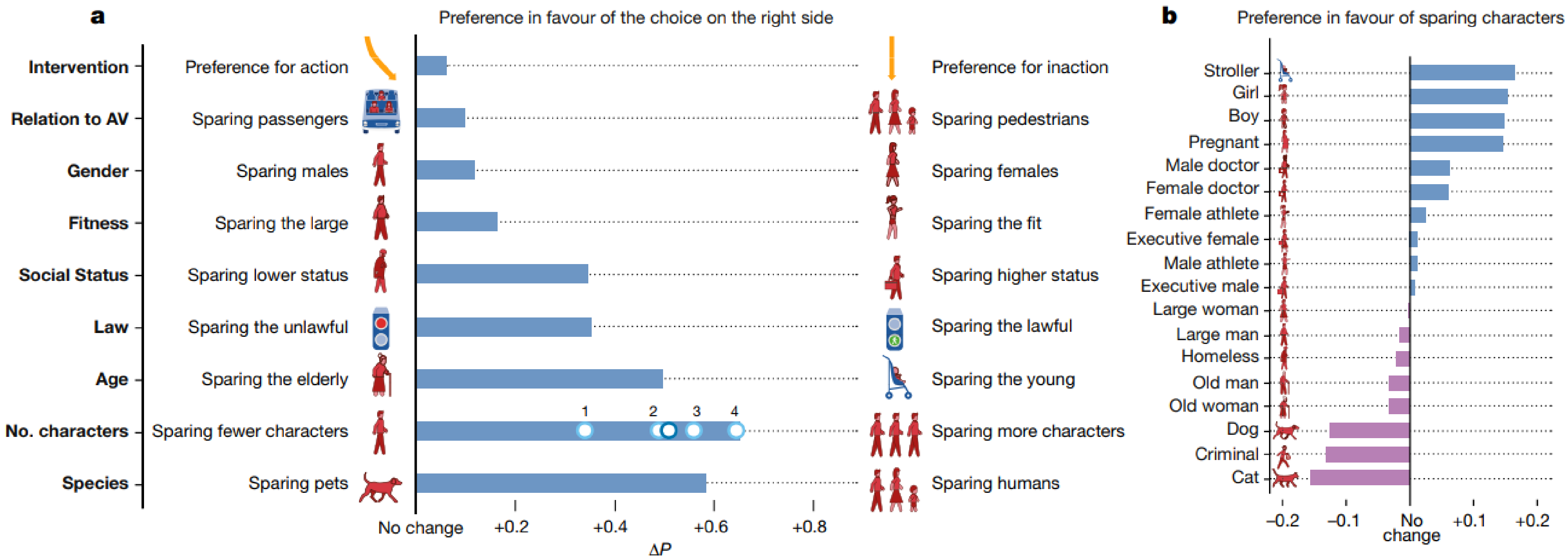


Fig. 2 | Global preferences. a, AMCE for each preference. In each row, ΔP is the difference between the probability of sparing characters possessing the attribute on the right, and the probability of sparing characters possessing the attribute on the left, aggregated over all other attributes. For example, for the attribute age, the probability of sparing young characters is 0.49 (s.e. = 0.0008) greater than the probability of sparing older characters. The 95% confidence intervals of the means are omitted owing to their insignificant width, given the sample size ($n = 35.2$ million). For the number of characters (No. characters), effect sizes are shown for each number of additional characters (1 to 4; $n_1 = 1.52$ million, $n_2 = 1.52$ million, $n_3 = 1.52$ million, $n_4 = 1.53$ million); the effect size for two additional characters overlaps with the mean effect of the attribute. AV, autonomous vehicle. **b**, Relative advantage or penalty for each character, compared to an adult man or woman. For each character, ΔP is the difference the between the probability of sparing this character (when presented alone) and the probability of sparing one adult man or woman ($n = 1$ million). For example, the probability of sparing a girl is 0.15 (s.e. = 0.003) higher than the probability of sparing an adult man or woman.

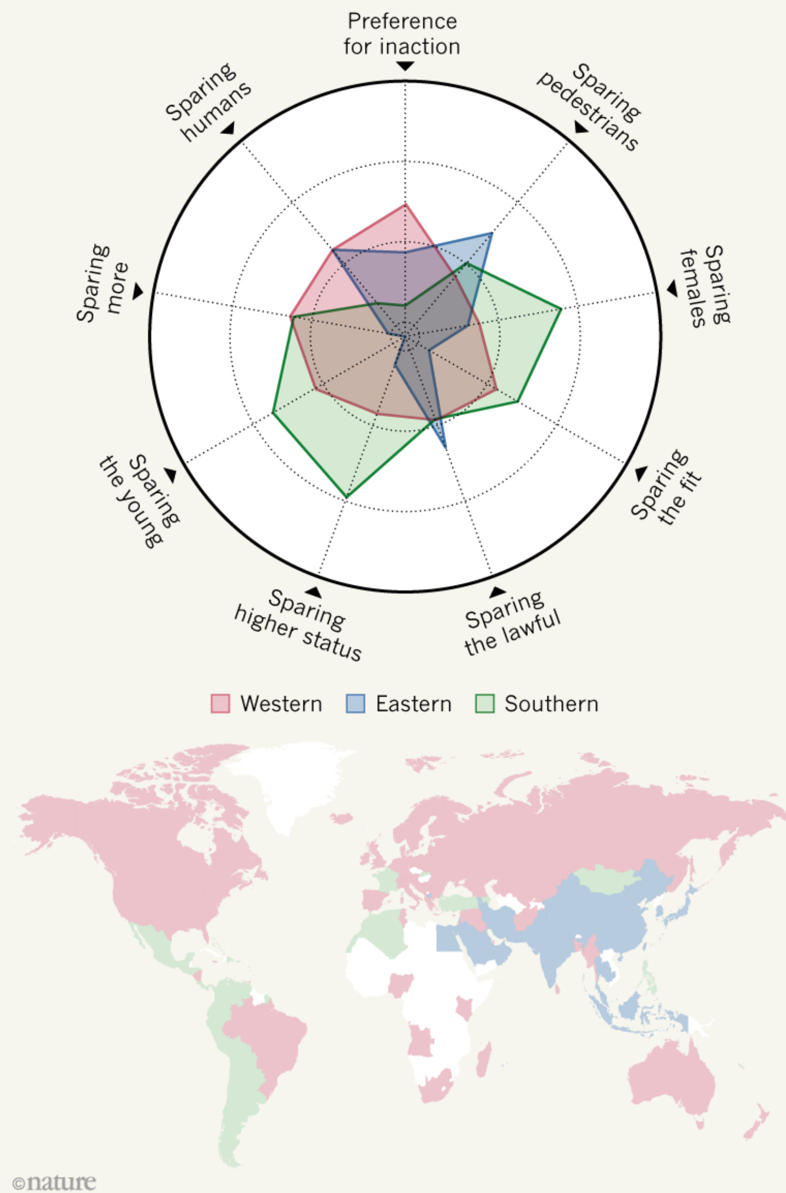
- AMCE (Average Marginal Component Effect)

Cultural Clusters

- Three large clusters
  - Western: Protestant, Catholic, and Orthodox countries in Europe and North America
  - Eastern: Islamic and Confucian (Asian) cultures
  - Southern: Central and South America, as well as France and former French colonies.
- The preference to spare younger characters rather than older characters is much less pronounced for countries in the Eastern cluster, and much higher for countries in the Southern cluster.
- The same is true for the preference for sparing higher status characters.
- Countries in the Southern cluster exhibit a much weaker preference for sparing humans over pets, compared to the other two clusters.
- Only the (weak) preference for sparing pedestrians over passengers and the (moderate) preference for sparing the lawful over the unlawful appear to be shared to the same extent in all clusters

MORAL COMPASS

A survey of 2.3 million people worldwide reveals variations in the moral principles that guide drivers' decisions. Respondents were presented with 13 scenarios, in which a collision that killed some combination of passengers and pedestrians was unavoidable, and asked to decide who they would spare. Scientists used these data to group countries and territories into three groups based on their moral attitudes.



AV Ethical Issues: is it Worth the Time?

- Many argue that ethical issues are just a distraction from the real problem of AV safety and security, esp. in the presence of ML/DL algorithms.
  - None of the AV accidents in recent years involved any ethical decisions similar to the Trolley Problem. They are due to failures in sensors or perception algorithms.
- Sebastian Thrun (formal head of Google’s SDC project, former professor at Stanford who led the development of Stanley, winner of DARPA Grant Challenge in 2005):
  - “I think it’s a great thing for philosophers to discuss these kind of problems. They can get tenure at their universities, but it's not of practical relevance. If we manage with certain car technology to halve the traffic deaths in the world, which means if we are able to have 500,000 fewer deaths in total, then for this extremely rare, purely hypothetical trolley problem that might occur once in a hundred years. I think whatever the outcome is, the mental energy that philosophers have spent on discussing it is completely out of proportion to the benefit of others on one problem. I will leave it at that.”

AV Testing Legislation (USA)

- It is absolutely necessary to test AVs on public roads for technology development, but is it ethical?
- Legislation regulating AV testing differs widely across states. Several states have no proposed legislation, meanwhile states like Nevada, California, Texas, and Arizona are hotbeds for testing AVs.
- My opinion: Yes, if Human Operators are alert and responsible.