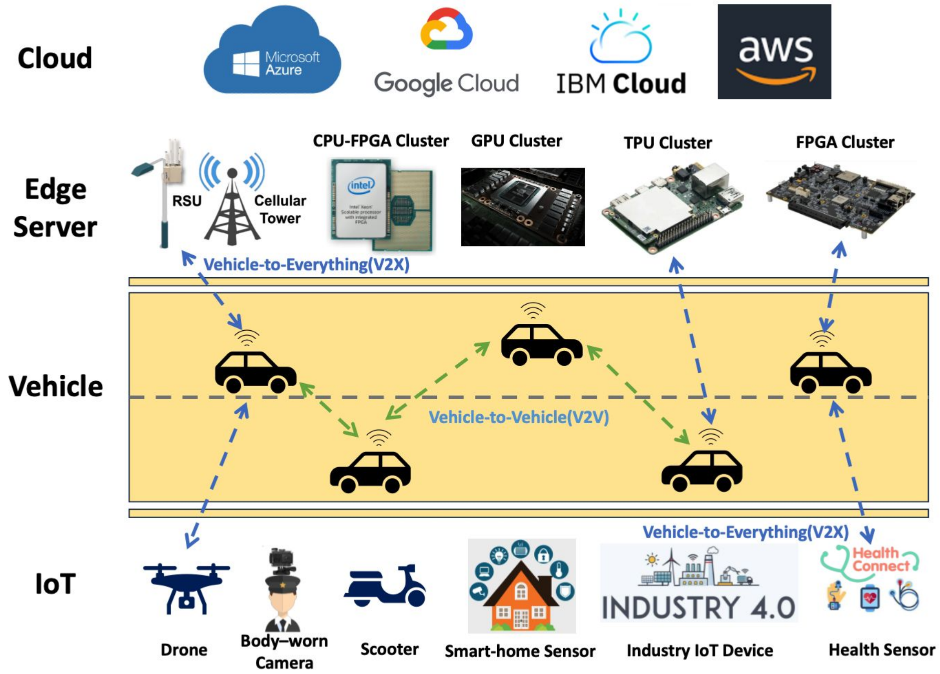
Quantitative Analysis of Storage Requirement for Autonomous Vehicles

Yuxin Wang, William He, Ruijun Wang and Weisong Shi

@The CAR Lab

July 8, 2024
HOTSTORAGE'24

Motivation



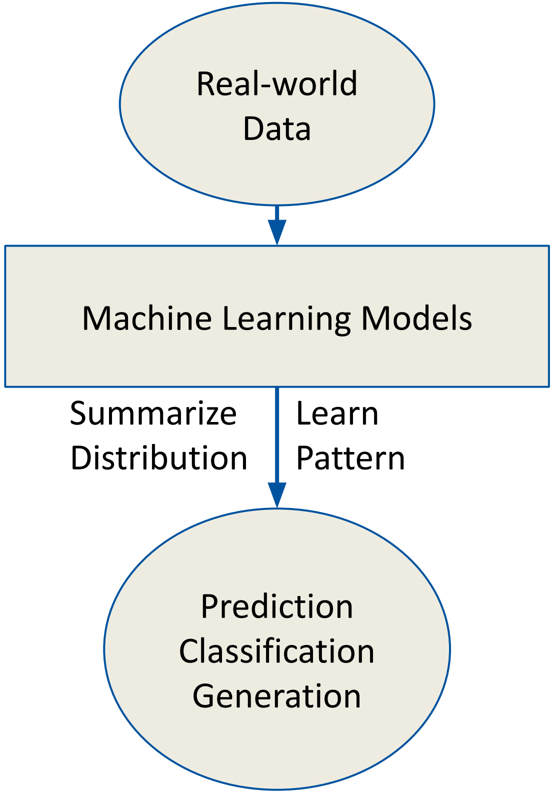


Image source: [Vehicle as a Mobile Computing Platform]

Motivation

Bermuda Image source: wikipedia

Motivation

1. Computation Platform:
   Long-term historical data
   e.g., AV Developers
   3rd-party applications

2. AVBlackBox:
   Short-term historical data
   e.g., Insurance company
   AV developers

1. **What and how much data will be generated by autonomous vehicles?**
2. **How to build storage system for autonomous vehicles?**
   a. **Write**
   b. Read
3. How to store them?

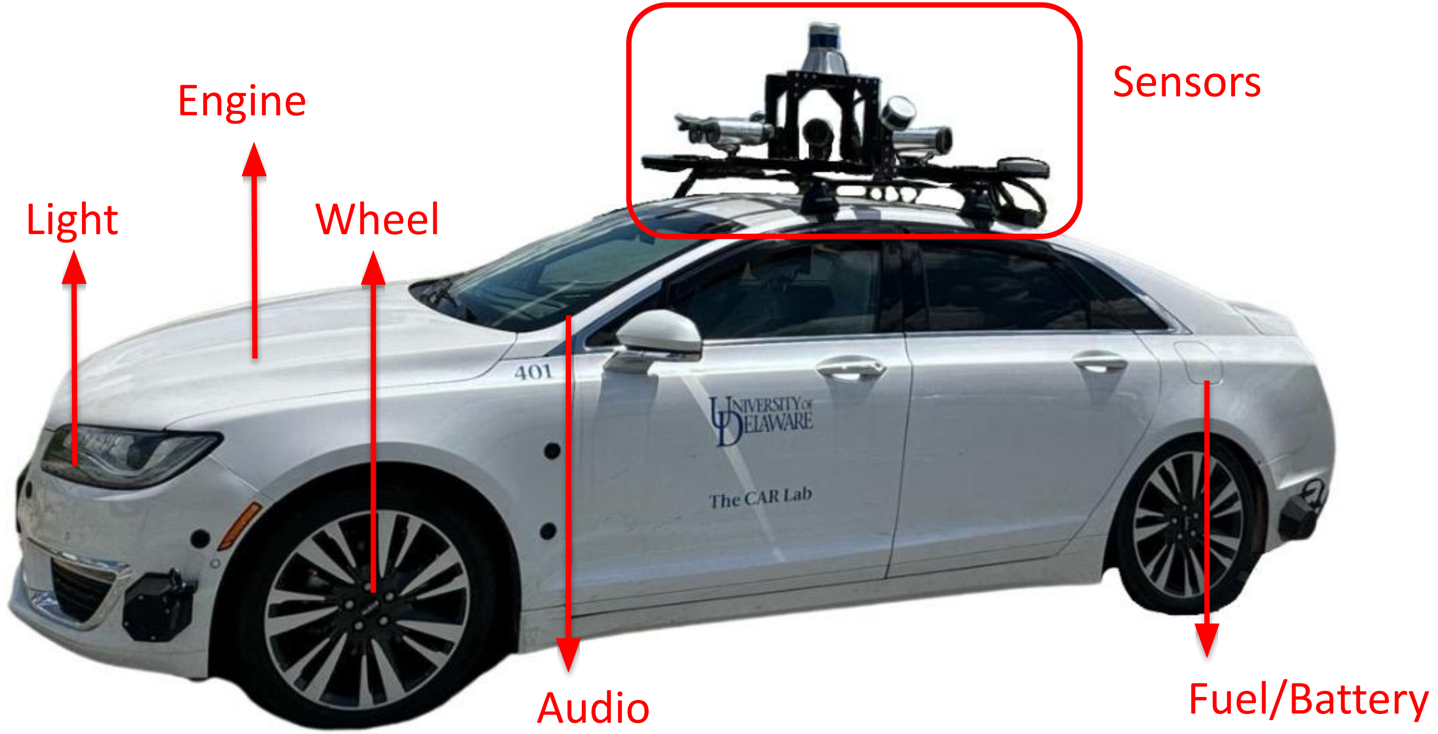
Our Contribution

- *Mathematical models* for sensor data rates
- Comprehensive evaluation in *multiple driving scenarios*, e.g., city, rural area, and highway
- A storage solution – AVStore
  - The detailed *storage requirement* based on using platform and sensor sets
  - Storage system *architecture* in autonomous vehicle

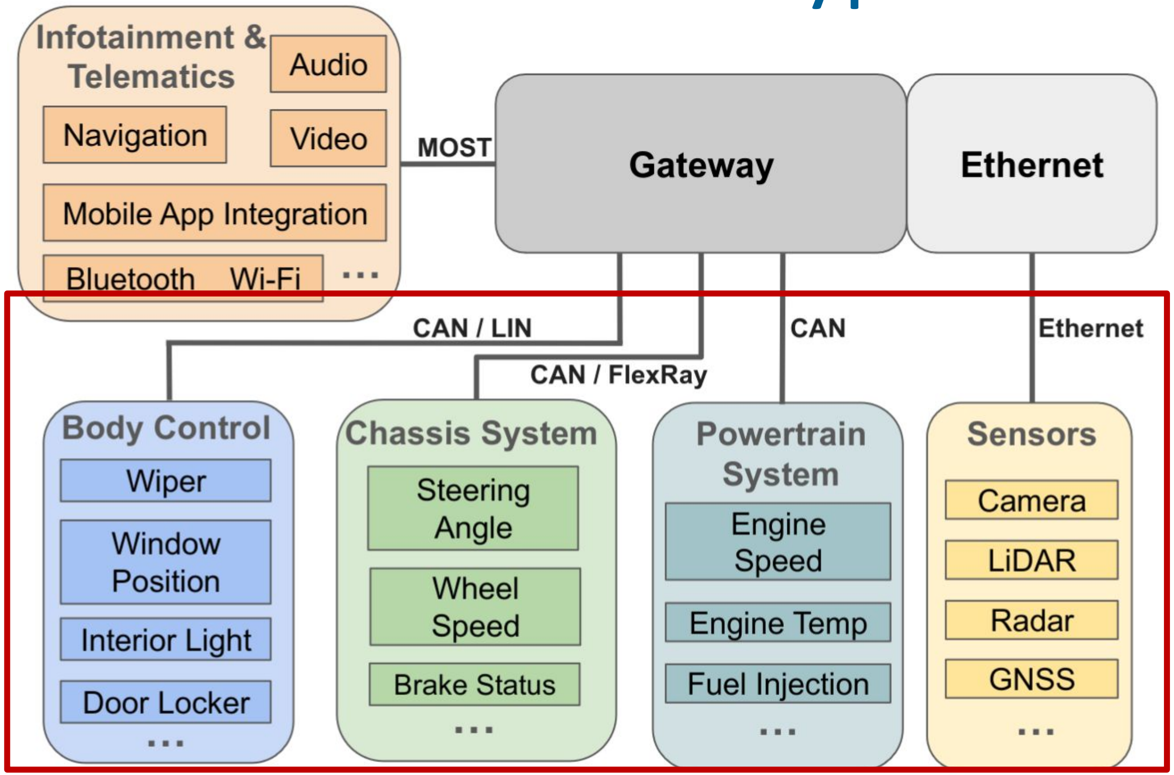
Outline

- Problem Statement
- Our Approach
- Evaluation Results
- Discussion

HydraU – An AV Platform



Data Type



Vehicle Status:
CAN bus data

Environment:
Sensor data

Sensor Data Model

- **CAN BUS Data**
  Maximum bit rate broadcast by each ECU is up to 1 million bits per second
- **RGB Camera**

$$DataRate_{RGB}(bytes/s) = \frac{1byte}{8bit} \times H \times V \times BPP \times FPS$$

- **3D LiDAR**

$$DataRate_{LiDAR}(bytes/s) = N \times B(bytes)$$

- **Radar**

$$DataRate_{Radar}(bytes/s) = f \times N_p \times B(bytes)$$

$$f = \frac{1}{CycleTime} \qquad N_p = \frac{FOV\ Azimuth}{Azimuth\ Resolution} \times \frac{FOV\ Elevation}{Elevation\ Resolution}$$

- **GNSS**

$$DataRate_{GNSS}(bytes/s) = r_{update} \times D(bytes)$$

All parameters could be found in the product user-manual or datasheet

Parameters Explanation

H: horizontal number of pixels

V: vertical number of pixels

BPP: pixel bit depth

FPS: frame rate per second

N: number of returned points

B: bit depth per point

f: scanning rate

N_p: the number of points that can be resolved within the radar's field of view (FOV)

r_{update}: the number of messages it generates per second

D: constant data size for each message

Camera Model

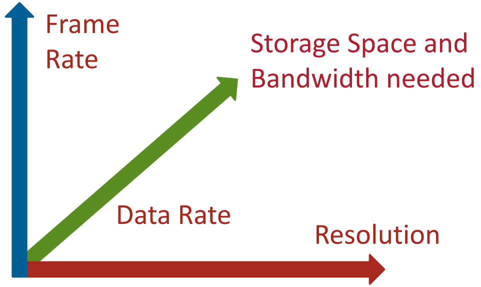
Data Rate(byte/s) = 1 byte / 8 bit × H × V × BPP × FPS

H: horizontal number of pixels

V: vertical number of pixels

BPP: pixel bit depth

FPS: frame rate per second



- Example1: *FL2-14S3M-C* used in KITTI (4 cameras)

Data Rate = 1384 x 1032 x 12 x 15 / 8

= ~ 32MB/s = ~*2.7TB/day*

- Example2: Basler ace used in our platform (7 cameras)

Data Rate = 1920 x 1200 x 8 x 5 / 8

= 11.52MB/s = ~*0.995TB/day*

Resolution	1384 x 1032
Frame Rate	15 FPS
Megapixels	1.4 MP
Chroma	Mono
Sensor Name	Sony ICX267
Sensor Type	CCD
Readout Method	Global shutter
Sensor Format	1/2"
Pixel Size	4.65 µm
Lens Mount	C-mount
ADC	12-bit

CAN Bus Data

- **CAN BUS Data**
  A controller area network (CAN) is a vehicle bus standard designed to enable efficient communication primarily between electronic control units (ECUs)
  Classical CAN baud rate: up to 1 million bits per second between each ECU

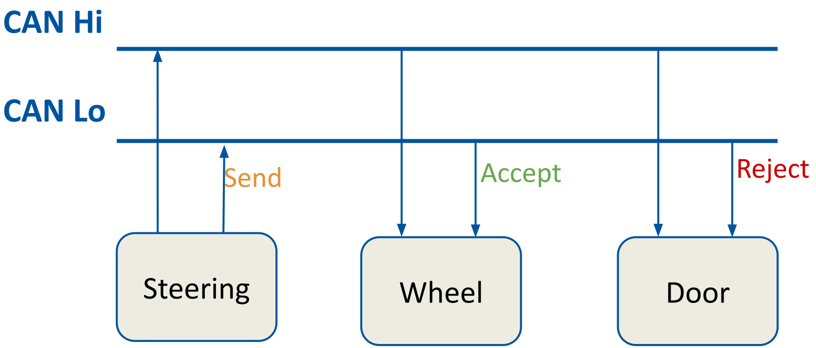


Fig.1 CAN bus

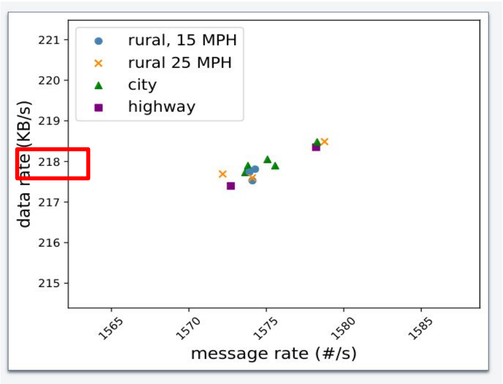


Fig.2 Experiment result

LiDAR Model

Data Rate (bytes/s) = N × B

N: number of returned points   B: bit depth per point

- **Data Format**

The **LAS** (LIDAR Aerial Survey) file format is a widely used binary file format designed to store 3D point cloud data collected by LiDAR.

- Example*(Hesai Pandar 64*, dual return mode

Data Rate = 2304000 × 26 = 59.904 MB/s

Data Points Generated	Single Return:	1,152,000 points/sec
	Dual Return:	2,304,000 points/sec

```
int offset = 0;
offset = addPointField(ros_msg, "x", 1, sensor_msgs::msg::PointField::FLOAT32, offset);
offset = addPointField(ros_msg, "y", 1, sensor_msgs::msg::PointField::FLOAT32, offset);
offset = addPointField(ros_msg, "z", 1, sensor_msgs::msg::PointField::FLOAT32, offset);
offset = addPointField(ros_msg, "intensity", 1, sensor_msgs::msg::PointField::FLOAT32, offset);
offset = addPointField(ros_msg, "ring", 1, sensor_msgs::msg::PointField::UINT16, offset);
offset = addPointField(ros_msg, "timestamp", 1, sensor_msgs::msg::PointField::FLOAT64, offset);
```

GPS Model

Data Rate(bytes/s) = r_{update} × D (bytes)

r_{update}: the number of messages it generates per second

D: constant data size for each message

- **Data Format**: GPS NMEA 0183 Sentences, **maximum** length of each sentences is 82 characters/bytes (data bits=8)
- Example (OEM7 GNSS kit)

  Topic: novatel/oem7/gps

  Frequency: 50Hz

  D: 394 bytes

  Data Rate = 50 * 394 = 19.7 KB/s

common header: 6
Oem7MessgeShortHeaderMem: 12
BESTGNSSPOSMem: 72
BESTPOSMem: 72
BESTVELMem: 44
PPPPOSMem:72
TERRASTARINFOMem:52
TERRASTARSTATUSMem:20
TIMEMem:44

Radar Model

Data Rate(byte/s) = $f \times N_p \times B$ (bytes)

$f = 1 /$ CycleTime

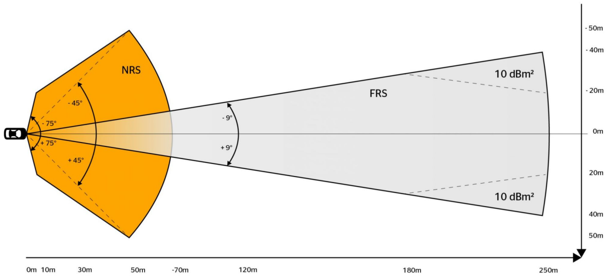
N_p = FOV Azimuth / Azimuth Resolution × FOV Elevation / Elevation Resolution

f: scanning rate

N_p: the number of points that can be resolved within the radar's field of view (FOV)

B: bit depth per point

- Example: *ARS430* record XYZ and density, point_step=16 bytes



FOV Azimuth	[°]	−9 … +9 (FRS) −60 … +60 (NRS)
Azimuth Beam Width (3dB)	[°]	2.2 (FRS)
Azimuth Accuracy	[°]	0.1 (FRS)
Azimuth Resolution	[°]	1.6 (FRS)
Antennas	–	12
Sensitivity (Min. RCS @ X M)	[m²]	10 @ 200m
Cycle Time	[ms]	75

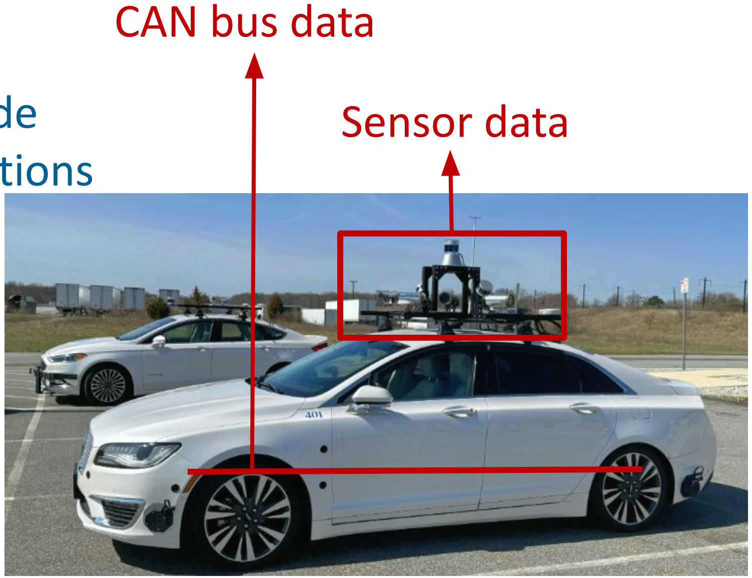
Outline

- Problem Statement
- Our Approach
- Evaluation Results
- Discussion

Experiment Design

- AV Platform:
  - 2018 Lincoln MKZ
  - CAN bus: drive-by-wire system
  - LiDAR: 1 Hesai Pandar64 on top, 2 VLP-16 on side
  - Camera: 7 Basler ace facing with different directions
  - GNSS/GPS: 2 Novatel OEM7

- Evaluation scenarios:
  - rural area with 15 miles per hour (MPH)
  - rural area with 25 miles per hour (MPH)
  - urban with varying speeds
  - highway with varying speeds



Experiment Result - Sensors

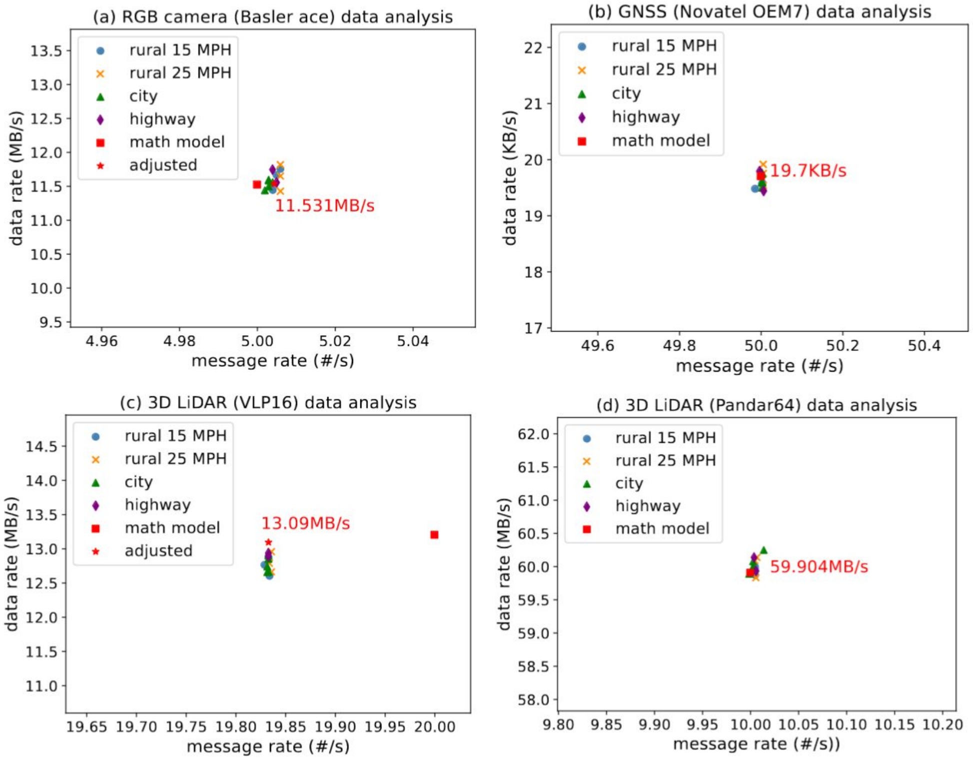


Fig. 1 : Data rate for each sensor.

Table 1 : Error for each sensor model.

Sensor	Error Mean	Error Max
RGB Camera	0.962%	2.454%
GNSS	0.699%	1.319%
VLP 16	2.230%	3.734%
Pandar 64	0.203%	0.569%

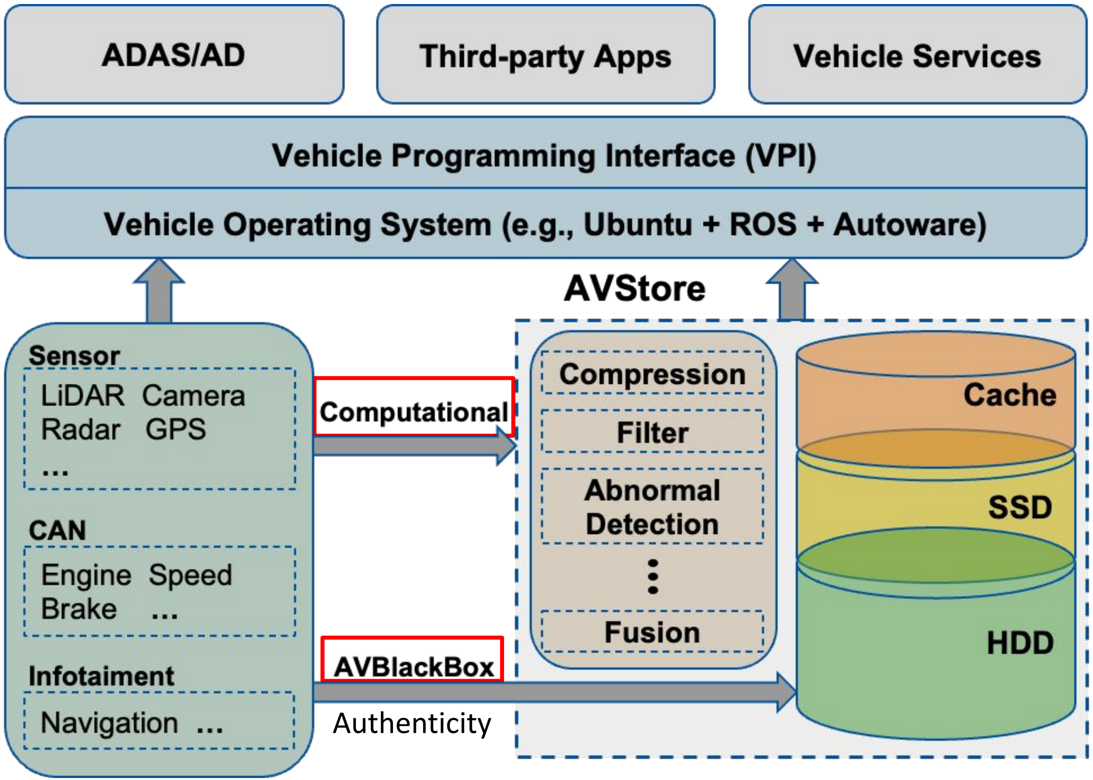
Storage Requirement

Type	Memory Access Requirements	Storage Size (collected)	Storage Size (math model)
RGB Camera (Basler ace)	11.530MB/s	1.001TB/day	0.996TB/day
3D LiDAR (VLP16)	12.788MB/s	1.105TB/day	1.131TB/day
3D LiDAR (Pandar64)	60.009MB/s	5.185TB/day	5.176TB/day
GNSS (Novatel OEM7)	19.624KB/s	1.696GB/day	1.702GB/day
CAN	217.833KB/s	18.821GB/day	-
Total	166.915MB/s	14.424TB/day	14.432TB/day

How do we store this huge streaming data onboard in real-time?

Vehicle Usage:
Average Driving Time in the US([AAA 2023 report]): 60.2mins/day

AVStore Architecture



An Example for Computational Storage

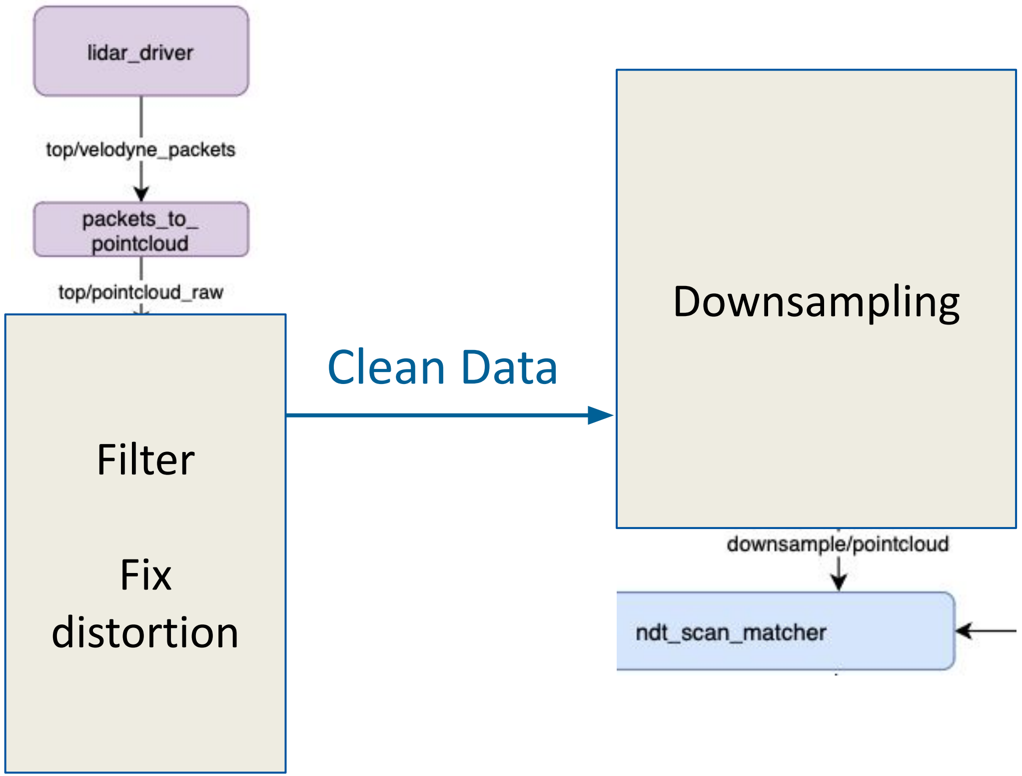
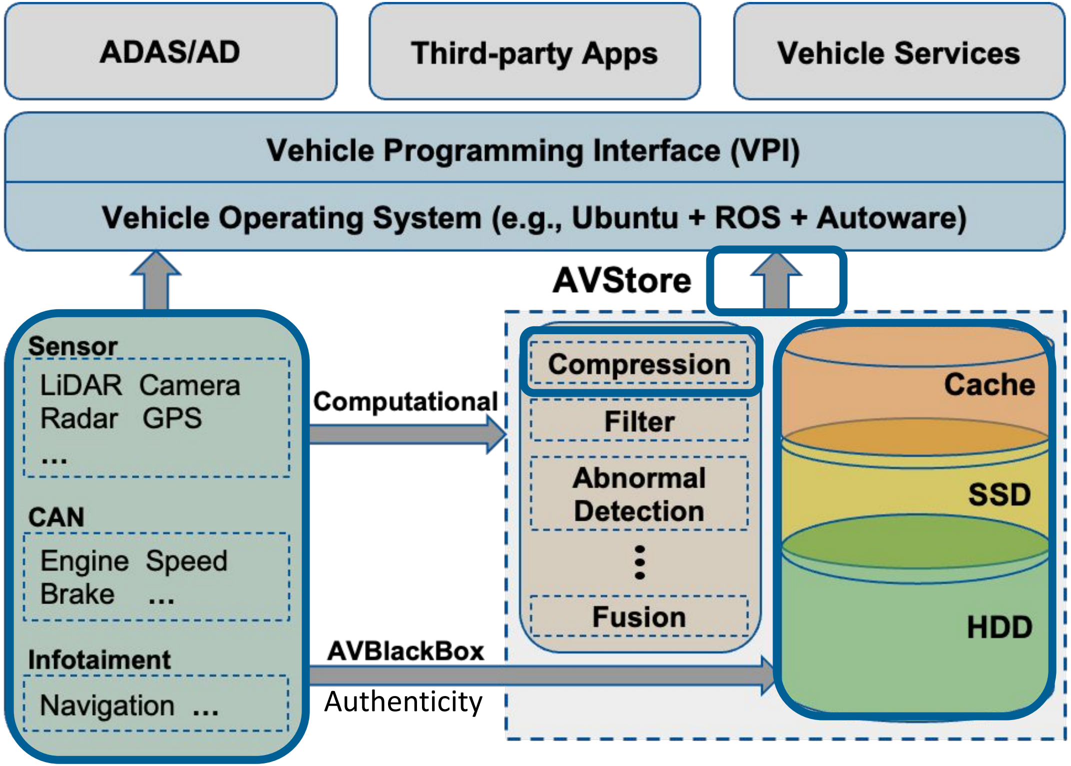


Image source: autoware

Outline

- Problem Statement
- Our Approach
- Evaluation Results
- Discussion

Discussion



? What data need to be collected?

? Data duplication?

? Data compression?

? What is the data using characteristics and retrieval frequency?

? Storing format?

? New structure?

Recap and Future Work

- **High Data Rate**: sensor data rate could reach tens to hundreds of MB/s (for one)
- **Context-agnostic:** CAN bus data rate and sensor data rate are not influenced by driving speed or the environment
- **Diversity**: different sensor has a very different data format
- **Opportunities**: the upper-level applications do not use all raw data directly context-aware sensing and storage

Thank you!

yuxw@udel.edu

?

http://thecarlab.org

Literature Review

Field	Publications	Related Part Summary
Industry	Flash Memory in the emerging age of autonomy. *Stephan Heinrich, 2017 (Lucid Motors)*	Some estimates on the raw sensor data generated by each type of sensor.
	Storage Capacity Requirement for Autonomous Vehicles to Balloon Over 2TB in the Next Decade. *Counterpoint Technology Market Research, 2019*	The amount of data generated by the growing sensor system at each level of autonomy, and the corresponding rise in the amount of in-vehicle storage needed.
Academic	HydraSpace: Computational data storage for autonomous vehicles. *R. Wang et al. 2020 (doi)*	Proposed multi-layered storage architecture and practical compression algorithms to manage the sensor pipe data.
	Cybersecurity and Capacity Requirement for Data Storage of Autonomous Driving System. *I. Kim et al. 2022 (doi)*	Estimate storage requirement based on rough data bandwidth used from Lucid Motors Report.
	Data Storage System Requirement for Autonomous Vehicle. *I. Kim et al. 2022 (doi)*	Roughly estimate data storage capacity for L3 and L4 based on open autonomou vehicles dataset.

Appendix

Appendix Table 1: CAN BUS data

scenarios	Duration(s)	Size(MB)	# Message	Data Rate(KB/s)	Message Rate(#/s)
Rural 15 MPH	282	61.4	443849	217.730	1573.933
	264	57.5	415614	217.803	1574.295
	274	59.6	431310	217.518	1574.124
Rural 25 MPH	199	43.3	313250	217.588	1574.121
	181	39.4	284570	217.680	1572.210
	184	40.2	290496	218.478	1578.783
City	314	68.6	495588	218.471	1578.306
	395	86.0	621594	217.722	1573.656
	521	113.6	820624	218.042	1575.094
	369	80.4	581395	217.886	1575.596
	246	53.6	387164	217.886	1573.837
Highway	229	50.0	361414	218.341	1578.227
	230	50.0	361728	217.391	1572.730

Appendix

Appendix Table 2: RGB camera (Balser ace) data

scenarios	Duration(s)	Size(GiB)	# Message	Data Rate(MB/s)	Message Rate(#/s)
Rural 15 MPH	253.389	2.7	1268	11.441	5.004
	220.997	2.4	1106	11.661	5.005
	246.697	2.7	1235	11.752	5.006
Rural 25 MPH	163.593	1.8	819	11.814	5.006
	159.794	1.7	800	11.423	5.006
	184.396	2.0	923	11.646	5.006
City	333.797	3.6	1670	11.580	5.003
	379.797	4.1	1900	11.591	5.003
	516.400	5.5	2583	11.436	5.002
	364.395	3.9	1823	11.492	5.003
	241.597	2.6	1209	11.555	5.004
Highway	237.794	2.6	1190	11.740	5.004
	278.896	3.0	1396	11.550	5.005

Appendix

Appendix Table 3: 3D LiDAR (VLP16) data

scenarios	Duration(s)	Size(GiB)	# Message	Data Rate(MB/s)	Message Rate(#/s)
Rural 15 MPH	258.462	3.1	5126	12.878	19.832
	218.769	2.6	4338	12.761	19.829
	247.106	2.9	4901	12.601	19.834
Rural 25 MPH	161.174	1.9	3197	12.658	19.836
	157.493	1.9	3124	12.954	19.836
	184.787	2.2	3665	12.784	19.834
City	334.296	4.0	6630	12.847	19.833
	387.76	4.6	7690	12.737	19.832
	525.977	6.2	10431	12.656	19.832
	373.163	4.4	7401	12.661	19.833
	249.781	3.0	4954	12.896	19.833
Highway	224.974	2.7	4462	12.886	19.833
	249.125	3.0	4941	12.930	19.833

Appendix

Appendix Table 4: 3D LiDAR (Hesai Pandar64) data

scenarios	Duration(s)	Size(GiB)	# Message	Data Rate(MB/s)	Message Rate(#/s)
Rural 15 MPH	225.987	12.6	2261	59.867	10.005
	238.188	13.3	2383	59.956	10.005
	230.892	12.9	2310	59.990	10.005
Rural 25 MPH	207.305	11.6	2074	60.083	10.005
	149.995	8.4	1501	60.132	10.007
	209.983	11.7	2101	59.827	10.006
City	334.895	18.7	3350	59.956	10.003
	384.288	21.5	3844	60.073	10.003
	521.770	29.1	5217	59.884	9.999
	368.695	20.6	3688	59.993	10.003
	245.958	13.8	2463	60.245	10.014
Highway	273.202	15.3	2733	60.132	10.004
	245.462	13.7	2456	59.929	10.006

Appendix

Appendix Table 5: GNSS (Novatel OEM7) data

scenarios	Duration(s)	Size(MiB)	# Message	Data Rate(KB/s)	Message Rate(#/s)
Rural 15 MPH	262.498	4.9	13126	19.574	50.004
	220.780	4.1	11036	19.473	49.986
	246.579	4.6	12330	19.561	50.004
Rural 25 MPH	166.421	3.1	8322	19.532	50.006
	159.301	3.0	7966	19.747	50.006
	184.339	3.5	9218	19.909	50.006
City	334.198	6.3	16711	19.766	50.003
	390.617	7.3	19532	19.596	50.003
	535.417	10.0	26773	19.584	50.004
	377.198	7.1	18861	19.737	50.003
	252.798	4.7	12641	19.495	50.004
Highway	227.895	4.3	11394	19.785	49.997
	140.241	2.6	7013	19.440	50.007