THE DEBS 2019 Grand Challenge

Oleh Bodunov, Vincenzo Gulisano, Hannaneh Najdataei,
Zbigniew Jerzak, André Martin, Pavel Smirnov,
Martin Strohbach, Holger Ziekow

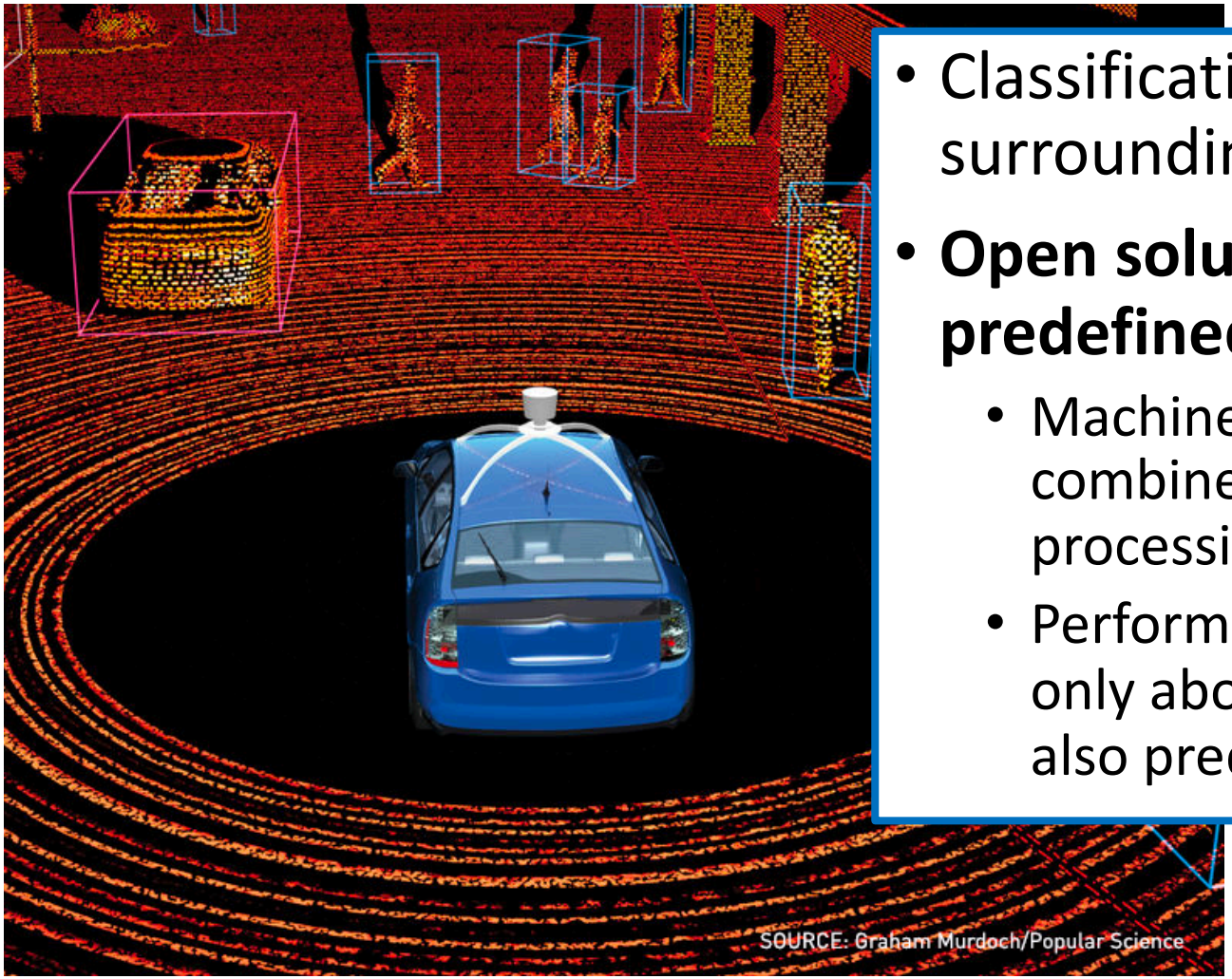
The DEBS Grand Challenge

- Started in 2011 (9th Grand Challenge in 2019)
- Provides a common ground and uniform evaluation criteria for a competition aimed at both research and industrial event-based systems.

Agenda

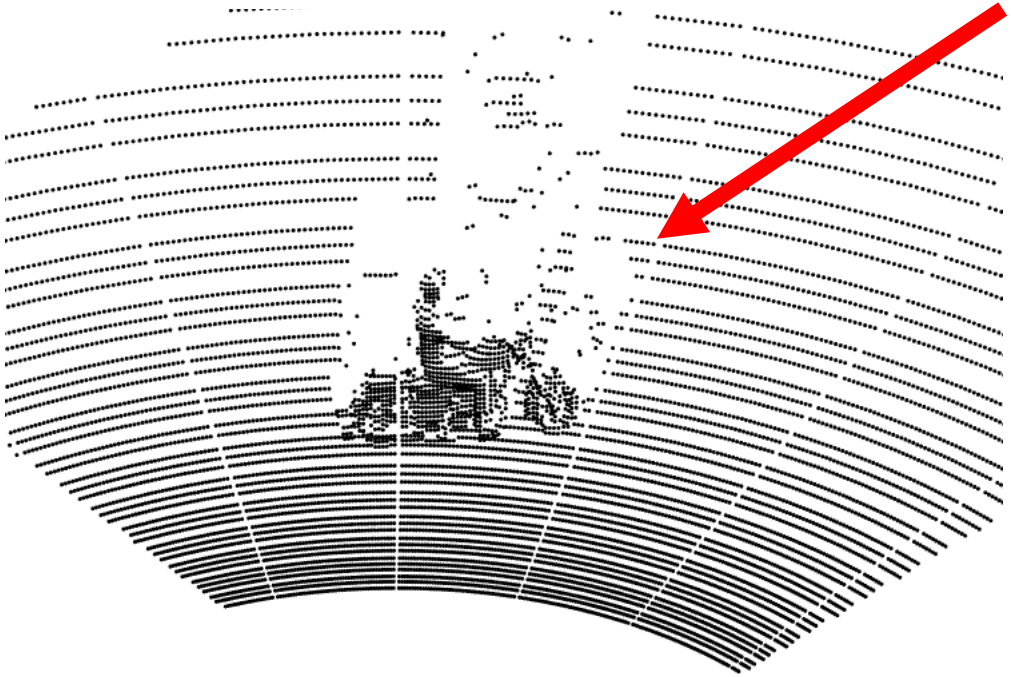
- This year’s challenge
- Evaluation Process
- How this final works

Sensing the environment with 3D laser scanners



- Classification of surrounding objects
- **Open solution: no predefined algorithm**
  - Machine learning combined with stream processing
  - Performance Award not only about speed but also prediction quality

The Problem



The Data 1/2

source: https://velodynelidar.com/hdl-64e.html

- Simulated point cloud readings (Webots)
- LiDAR sensor:
  - 64 lasers
  - 1125 reading per rotation / per laser
  - 72,000 points per rotation
- Readings carry
  - ts, laser, X, Y, Z coordinates

source: https://velodynelidar.com/hdl-64e.html

The Data 2/2

- Objects found in scenes:
  - ATM machine
  - pedestrian
  - benches
  - cloth recycling container
  - drinking fountain

  …

~30 labels for objects found in urban environments

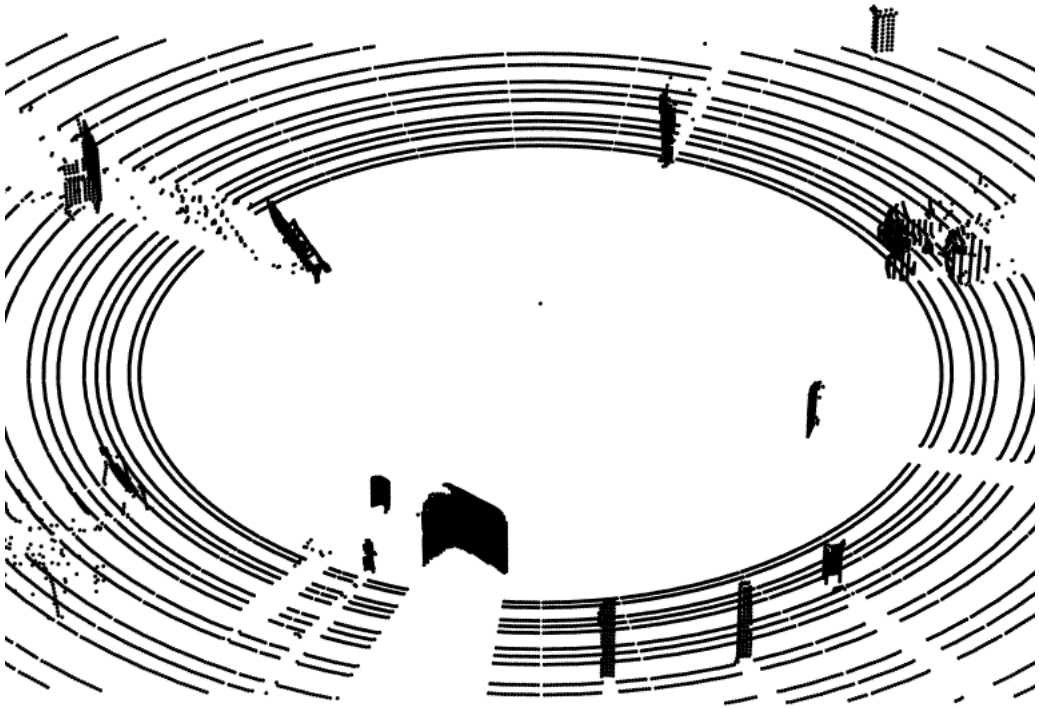
Problem Details

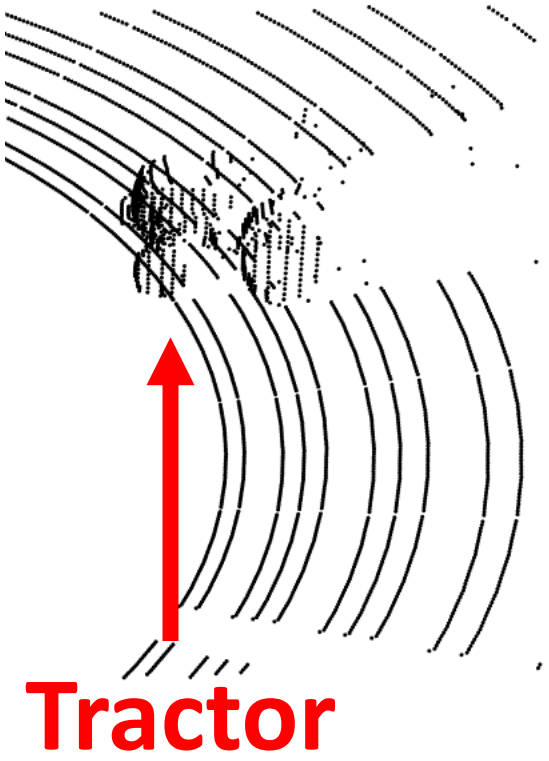
- Objects may be very close to each other, making it hard to separate the point clouds

- Objects may partially (or even totally) occlude other objects in the scene
  - due to total occlusion, a perfect solution of the challenge is not necessarily possible

- Even without occlusion by other objects, the back facing part of the object is always occluded

Sample scene





The overall Score

- Solutions are **ranked** by
  - (A) processing speed
    - Total runtime ($rank_0$)
    - Average per-scene latency ($rank_1$)
  - (B) classification quality
    - Accuracy ($rank_2$)
    - Precision ($rank_3$)
    - Recall ($rank_4$)

- Final rank is the sum of the ranks (lower = better)

The overall Score

$$Accuracy, A = \frac{1}{n}\sum_{i=1}^{n}\frac{|Y_i \cap Z_i|}{|Y_i \cup Z_i|}$$

ratio of correctly predicted objects in the scene to the total number of predicted and existing objects in the scene

Number of scenes

$$Precision, P = \frac{1}{n}\sum_{i=1}^{n}\frac{|Y_i \cap Z_i|}{|Z_i|}$$

ratio of correctly predicted objects to the total number of existing objects in the scene

$$Recall, R = \frac{1}{n}\sum_{i=1}^{n}\frac{|Y_i \cap Z_i|}{|Y_i|}$$

ratio of correctly predicted objects to the total number of predicted objects in the scene

HOBBIT Lite platform

TECHNISCHE UNIVERSITÄT DRESDEN

Benchmark customer

Requires ranking of alternative solutions by some KPI

1

Submit benchmarks

Solution provider (vendor)
(e.g. DB, Streaming Platforms, ML frameworks, etc…)

Submit systems

HOBBIT Lite platform
(online or local instance)

Provides:

1. Automated benchmark executions
2. Bulk executions (challenge)
3. GUI, Leaderboards

Advantages:

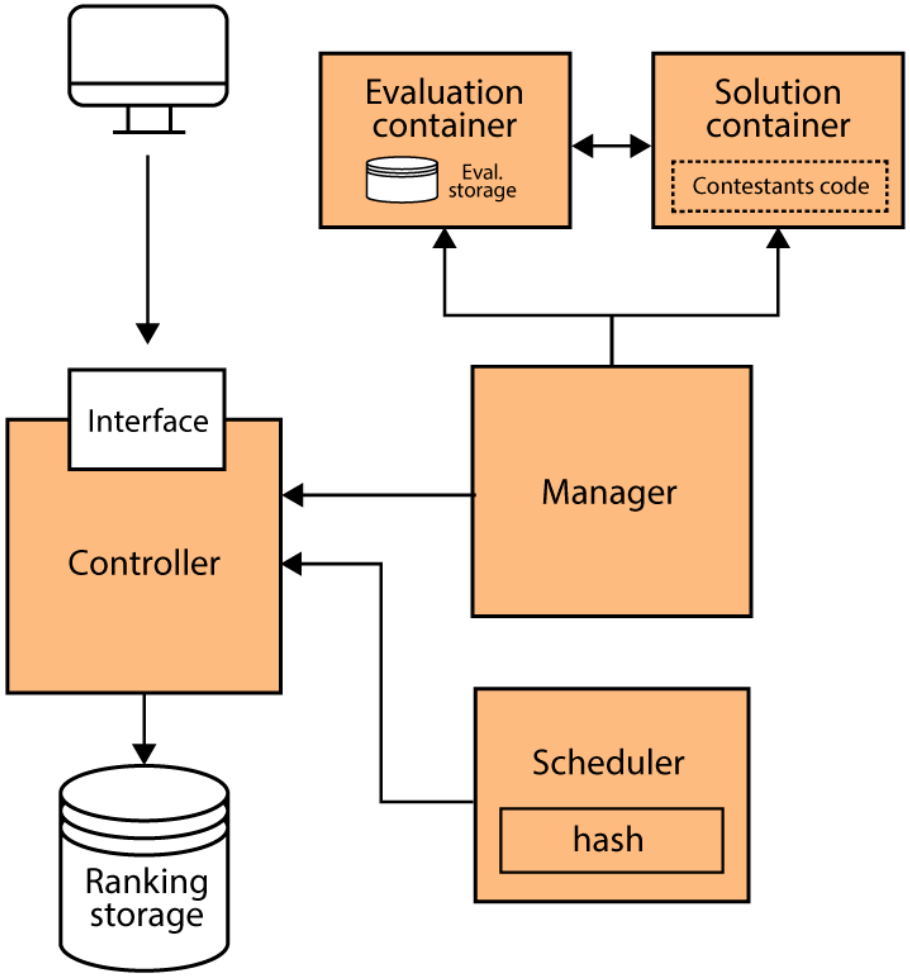
1. Batch/streaming benchmarks
2. Docker virtualization
3. Scalability
4. Ease of use

docker

docker hub

python

The HOBBIT Lite architecture



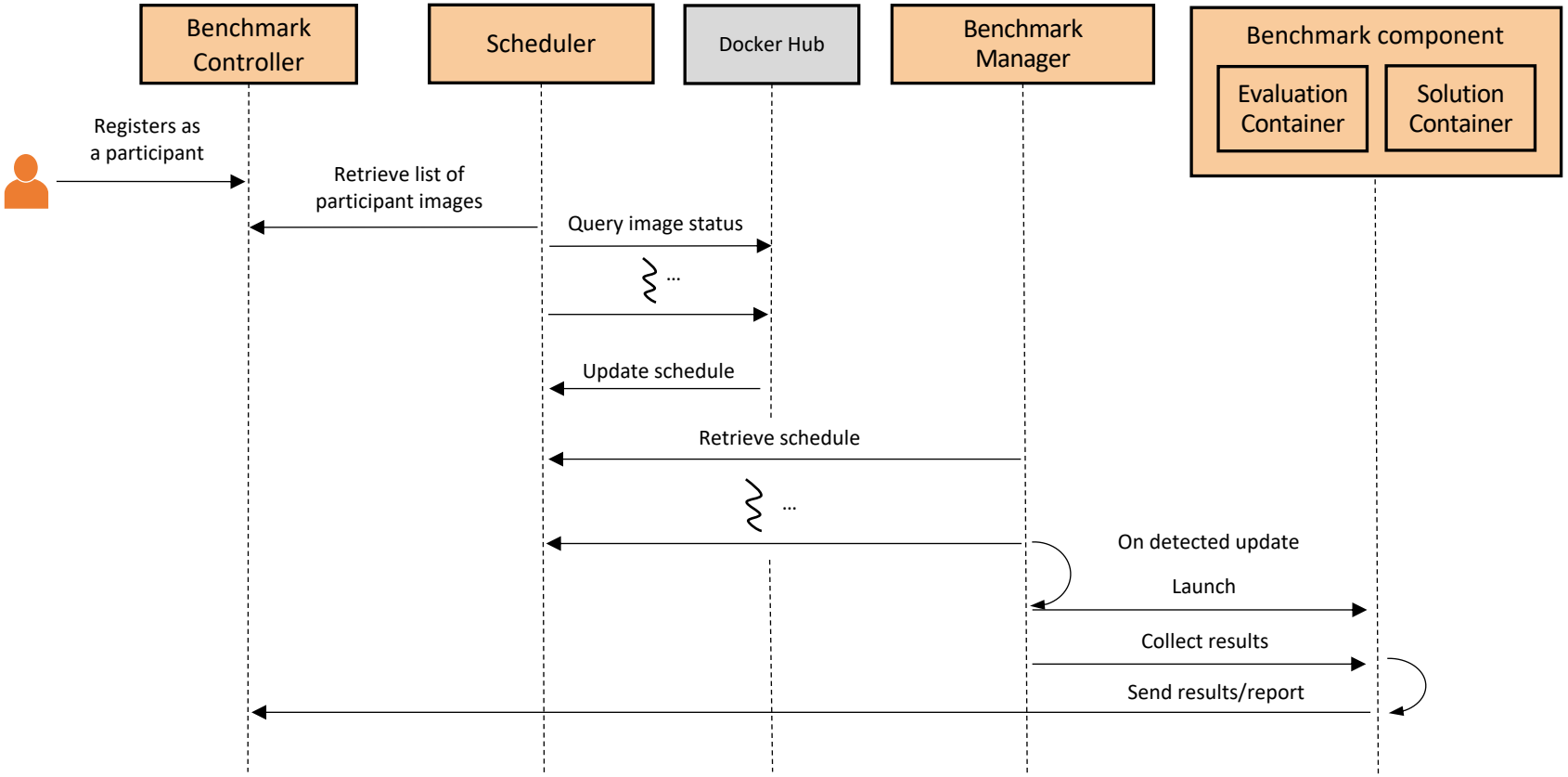
Platform specifications:

Cluster: 2 nodes.
Each 2x64 bit Intel Xeon E5-2683v3 (8-Cores, 2,0 GHz, Hyperthreading, 20MB Cache, each proc.), 116 GB RAM, 10Gb Ethernet

Written in **Python** and frameworks such as **Flask, Docker** and **Kubernetes**.

The DEBS GC 2019 Benchmark

TECHNISCHE UNIVERSITÄT DRESDEN



https://github.com/debs-2019-challenge/debs-2019-challenge

Dataset Statistics

- Initial dataset containing scenes with a single object per scene for all object types.
- Training and evaluation dataset containing scenes with multiple objects.
- All released datasets contain the input data and expected output
- Evaluation dataset:
  - 500 scenes
  - 10 to 80 objects per scene

2019 Statistics

6 Teams registered
in EasyChair

2 finalists
in this session

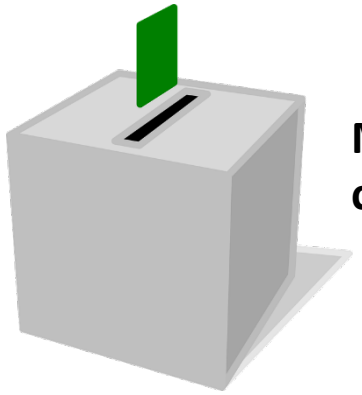
Awards

- **$1000 “Grand Challenge Award”**
  - for the best performing submission

Audience Awards (Voted for by you!)

- ***Most interesting/appealing solution***
- Voting boxes and valleys at the exits

Awards will be announced during the banquette



Make your vote count!

Results overview

- [rank 1] duration:20-25 minutes
- [rank 2] average-latency:2-3 seconds
- [rank 3] accuracy: 0.3-0.4
- [rank 4] precision: 0.4-0.5
- [rank 5] recall:0.6-0.7