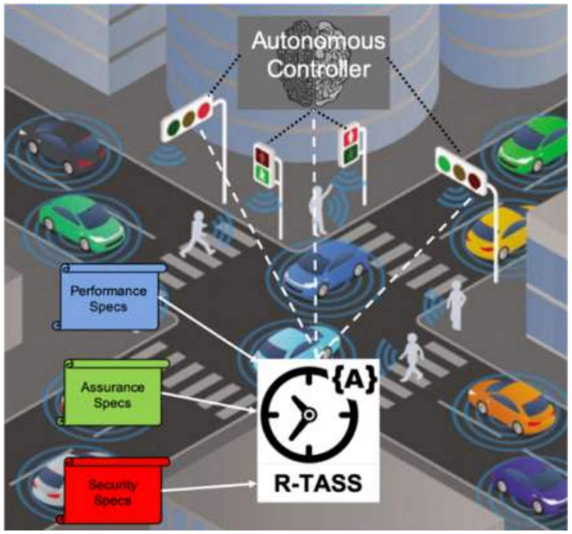
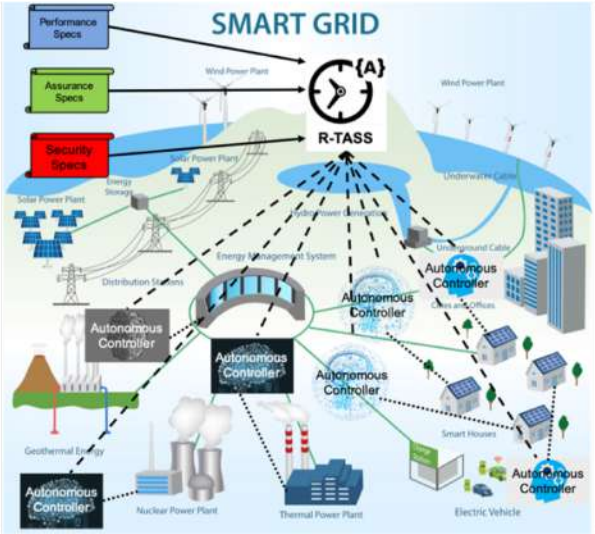
JOHNS HOPKINS
INSTITUTE *for*
ASSURED AUTONOMY

Assuring City Scale Infrastructure Systems

PI: Yair Amir, WSE
PI: Tamim Sookoor, APL
Total Budget: $750K (2 years)
03/16/2020

What Problem Are You Trying to Solve?





AI systems give better performance in the average case

They currently cannot be used in critical systems that need to guarantee the worst case

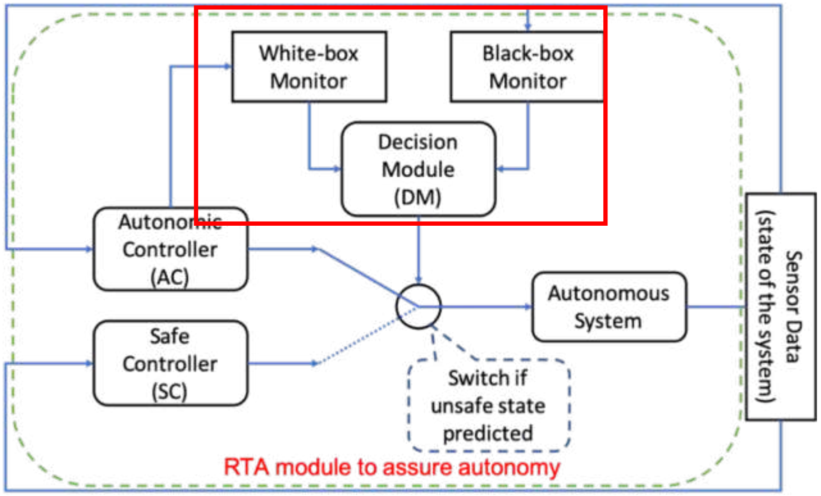
Assure the safety of autonomously controlled critical infrastructure systems

How is this done now? What's wrong with that?

- Currently AI systems are not used for critical infrastructure
- Deep learning systems have become very good at solving complicated problems. For many applications their expected performance is considerably better than alternative approaches
- Deep learning systems are hard to reason about and have non-intuitive failures on edge cases
- Current research believes that these edge cases are inherent and cannot be trained away
- The problem with AI systems is the edge cases. We want to gain the benefits of AI on the average case without paying for the cost of failures on the edge cases

What is new about your approach? Why do you think it might work? What are the risks?

- Designing a monitoring architecture that can take over from the AI with a safe controller when the AI system is at risk of breaching the correctness envelope
    - Adapting the standard approach for assured, dependable systems to the realm of assured AI
- Problem Risk: We won't be able to find a monitor and safe controller that is simple enough to be assured and still allows for good performance
- Meta Risk: We won't be able to generalize from the two problem areas we are looking at


White-box Monitor
Black-box Monitor
Decision Module (DM)
Autonomic Controller (AC)
Safe Controller (SC)
Autonomous System
Switch if unsafe state predicted
Sensor Data (state of the system)
RTA module to assure autonomy

Monitors

- Black Box Monitor
    - Looks at the state of the world and ensures safe behavior
- White Box Monitor
    - Looks at the state of the AI and ensures competent and confident decisions

Black Box Monitor

- The black box monitor looks at the state of the world and if the AI tries to put the system into a bad state, takes over and performs a simpler algorithm that has been provably assured to meet safety guarantees

- The black box is much simpler and does not change with training and thus can be reasoned about and proven correct from system invariants

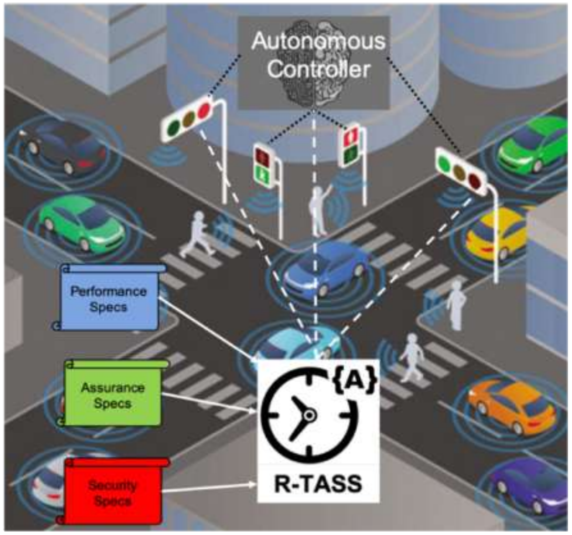
- The black box monitor allows us to prove the safety of the system since it only need to be proven about the black box system

White Box Monitor

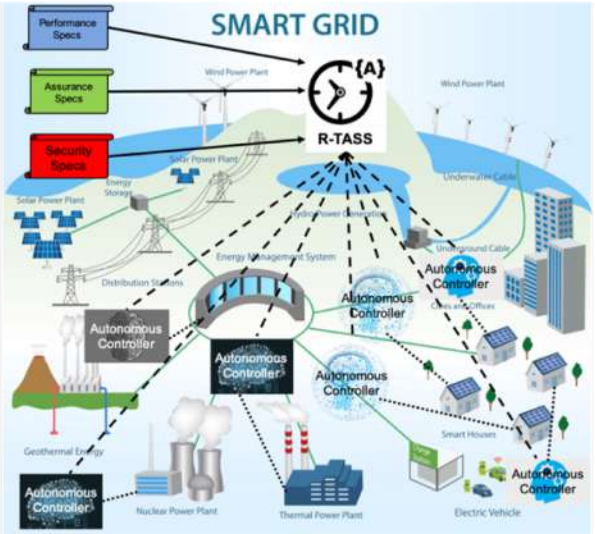
- The white box monitor looks at the state of the AI system
- Some examples include
  - How confident the AI is in its current decision
  - How similar the current input is to items that the system has trained on
  - What path is the AI expecting to do over the course of the next several time steps
- The white box monitor measures the certainty with which the AI system makes decisions as well as its sensitivity to noisy input data and feeds this information to the Decision Module which uses this along with information from the black box monitor to decide if control should be switched over to a simpler algorithm

Flagship Projects

- Implementing technology for runtime assurance of critical infrastructure systems
- Designing two ecosystem testbeds targeted at transportation and public safety domains



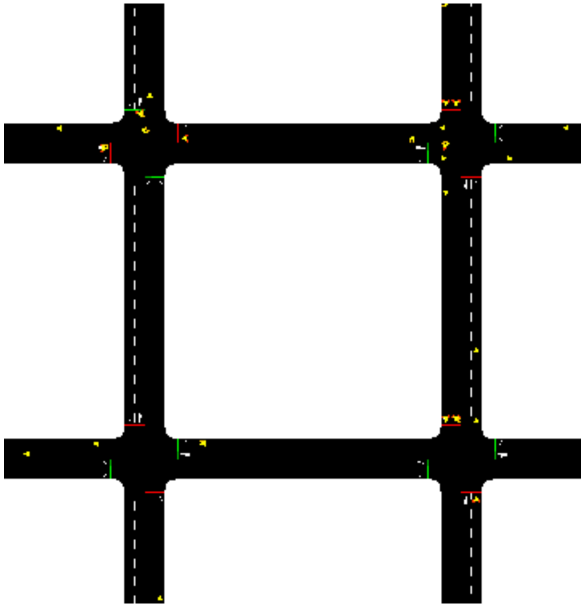
Transportation – Intelligent Traffic Control



Public Safety – Smart Power Grid

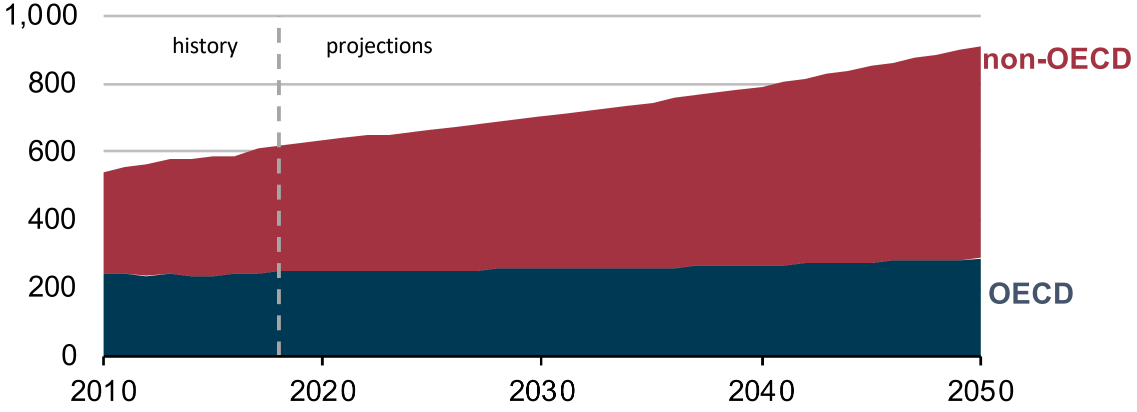
Testbeds could be used for transportation and public safety domains

Traffic Simulator Testbed Demo



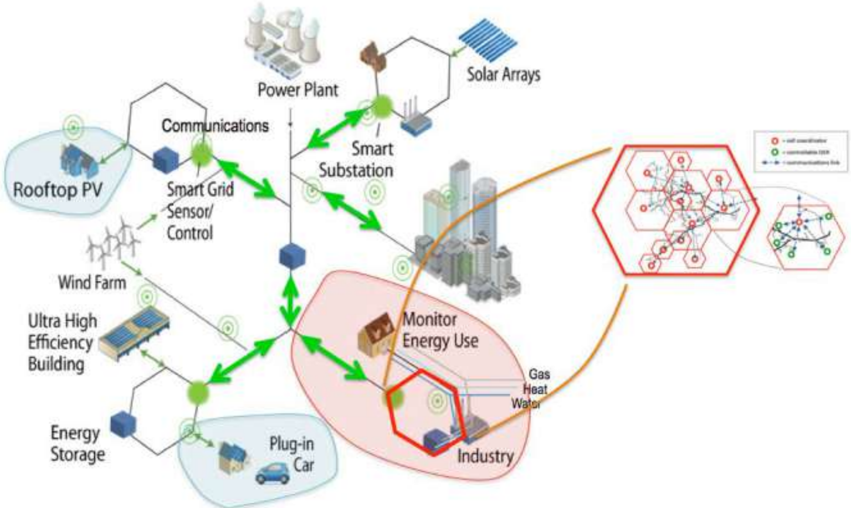
Enable autonomous energy grids

- **World energy consumption**
- quadrillion British thermal units



The *International Energy Outlook* 2019 (IEO2019)

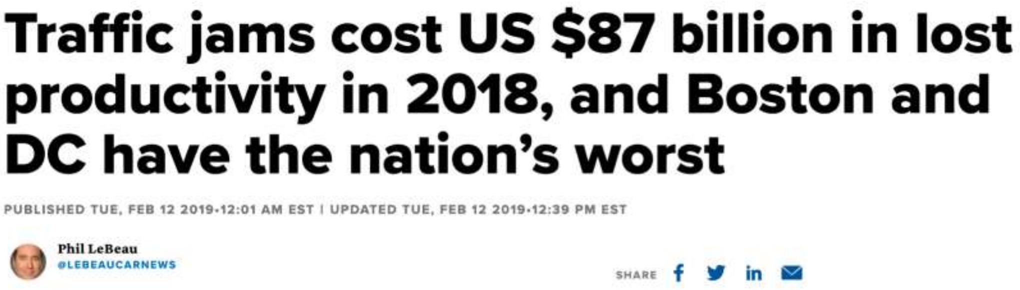
Autonomous energy grid
organized into self-optimizing cells



Kroposki, Benjamin D., et al. *Autonomous energy grids*. No. NREL/CP-5D00-68712. National Renewable Energy Lab.(NREL), Golden, CO (United States), 2017.

Autonomous grids necessary to support 50% consumption increase by 2050

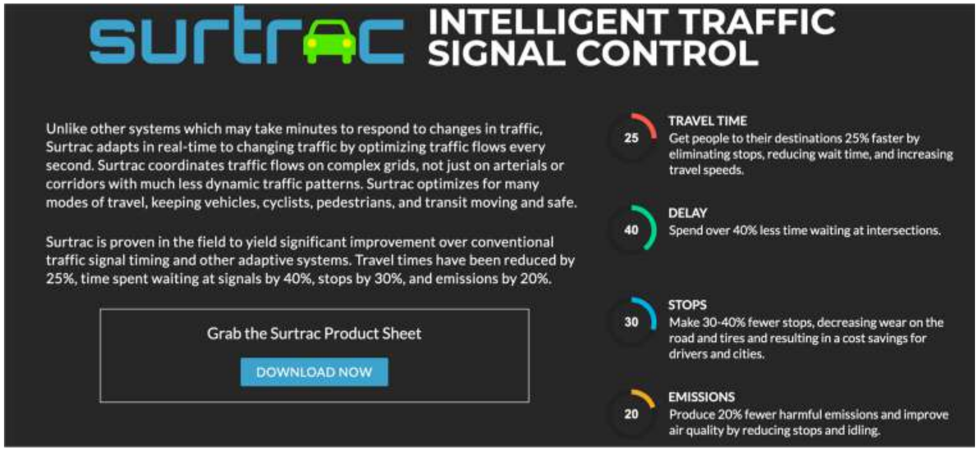
Widespread adoption of intelligent traffic control



Traffic jams cost US $87 billion in lost productivity in 2018, and Boston and DC have the nation's worst

PUBLISHED TUE, FEB 12 2019·12:01 AM EST | UPDATED TUE, FEB 12 2019·12:39 PM EST

Phil LeBeau
@LEBEAUCARNEWS



surtrac INTELLIGENT TRAFFIC SIGNAL CONTROL

Unlike other systems which may take minutes to respond to changes in traffic, Surtrac adapts in real-time to changing traffic by optimizing traffic flows every second. Surtrac coordinates traffic flows on complex grids, not just on arterials or corridors with much less dynamic traffic patterns. Surtrac optimizes for many modes of travel, keeping vehicles, cyclists, pedestrians, and transit moving and safe.

Surtrac is proven in the field to yield significant improvement over conventional traffic signal timing and other adaptive systems. Travel times have been reduced by 25%, time spent waiting at signals by 40%, stops by 30%, and emissions by 20%.

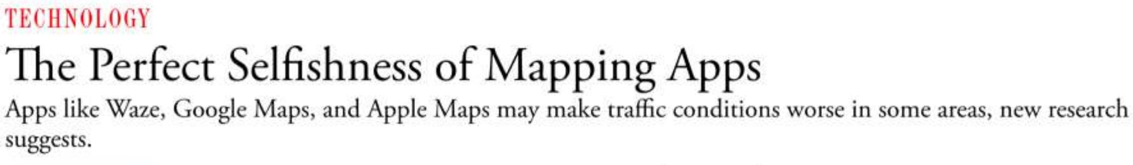
Grab the Surtrac Product Sheet

25 TRAVEL TIME
Get people to their destinations 25% faster by eliminating stops, reducing wait time, and increasing travel speeds.

40 DELAY
Spend over 40% less time waiting at intersections.

30 STOPS
Make 30-40% fewer stops, decreasing wear on the road and tires and resulting in a cost savings for drivers and cities.

20 EMISSIONS
Produce 20% fewer harmful emissions and improve air quality by reducing stops and idling.

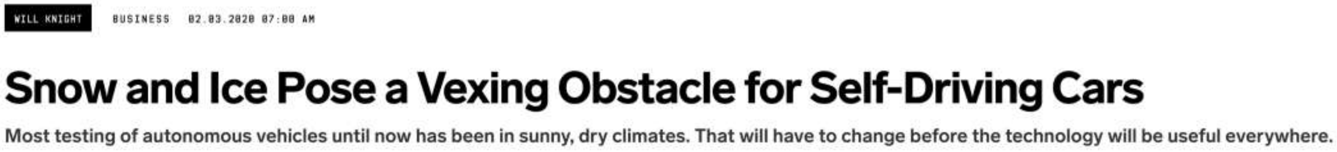


TECHNOLOGY

The Perfect Selfishness of Mapping Apps

Apps like Waze, Google Maps, and Apple Maps may make traffic conditions worse in some areas, new research suggests.

ALEXIS C. MADRIGAL MARCH 15, 2018



WILL KNIGHT BUSINESS 02.03.2020 07:00 AM

Snow and Ice Pose a Vexing Obstacle for Self-Driving Cars

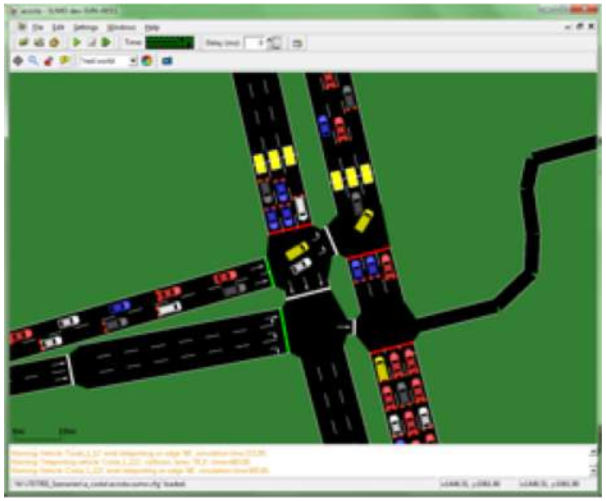
Most testing of autonomous vehicles until now has been in sunny, dry climates. That will have to change before the technology will be useful everywhere.

Reduce traffic jam costs by enabling safe traffic management systems

What will be accomplished in this project?

- Publications of our findings at DSN, ICDCS, SRDS, and/or TDSC on:
  - Extending Simplex to assure autonomous critical infrastructure systems
  - Runtime assurance of reinforcement learning-based autonomic controller
  - Challenges in assuring city-scale distributed systems
- RADICS: Runtime Assurance of Distributed Intelligent Control Systems will be demonstrated on:

Power Grid Testbed



Intelligent Traffic Control Testbed

Demonstrate runtime assurance on two testbeds by end of year 2

Anticipated External Funding: Energy Sector

- From which commercial and/or government sponsors to you anticipate seeking follow on funding?
  - We are getting a DOE grant to work on resilient power grids that includes both funding and contacts with three national labs (PNNL, LBNL, and Sandia), three utilities (HECO, PNM, WAPA), and three industry partners (ABB, GE, Siemens). We hope this will lead to continued funding along the lines of this proposal
- What levels of funding do you anticipate from which sponsor and on what timeline?
  - We do not know. The current project's total budget is about $5 million ($750k for DSN Lab)
- What help do you need from the IAA in pursing external funding?
  - We would appreciate any help with connections and exposure

Anticipated External Funding: Traffic Control

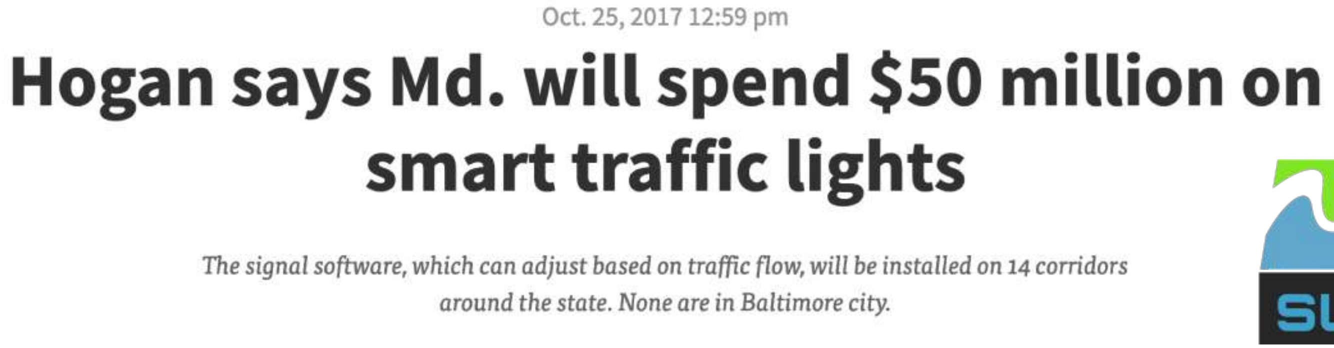
CATT
CENTER FOR ADVANCED
TRANSPORTATION TECHNOLOGY

UNIVERSITY OF MARYLAND 1856

U.S. Department
of Transportation
Federal Highway
Administration

MDOT
MARYLAND DEPARTMENT
OF TRANSPORTATION
STATE HIGHWAY
ADMINISTRATION

MDOT™
MARYLAND DEPARTMENT
OF TRANSPORTATION



Oct. 25, 2017 12:59 pm

Hogan says Md. will spend $50 million on smart traffic lights

The signal software, which can adjust based on traffic flow, will be installed on 14 corridors around the state. None are in Baltimore city.

Rapid Flow

surtrac INTELLIGENT TRAFFIC
SIGNAL CONTROL

IAA could engage with transportation administrations & technology vendors

Anticipated External Funding: Urban Air Mobility

- Existing and developing relationships with:

FEDERAL AVIATION ADMINISTRATION NextGEN

NASA

AIRBUS

AiRXOS Part of GE Aviation

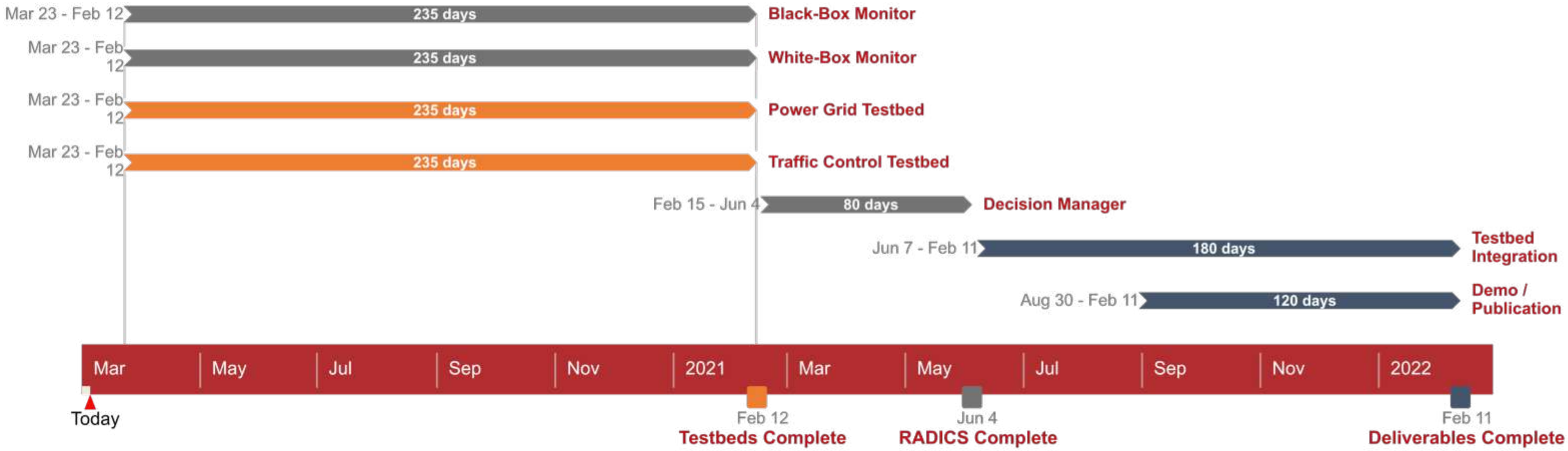
AIRMAP

UNIFLY®

Potentially millions of dollars of funding over the next decade

What is the cost and schedule?

	Year 1	Year 2
WSE	$120K	$122K
APL	$249K	$257K



Four parallel efforts in year 1 converge to a tech demo at the end of year 2

JOHNS HOPKINS

INSTITUTE *for*
ASSURED AUTONOMY