# MULTI-MODEL TESTBED FOR DEEP LEARNING-DRIVEN RESILIENT CPS

SOS LABLET MEETING -- 2019 JULY

HIMANSHU NEEMA, PHD
RESEARCH ASSISTANT PROFESSOR

PETER VOLGYESI
RESEARCH SCIENTIST

VANDERBILT UNIVERSITY

CPS SECURITY LABLET



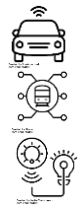
## TESTBED GOALS

**Repeatable experiments** with quantifiable metrics

**Integrated Machine Learning** components

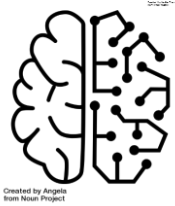
**Multiple CPS domains**

- Highway Transportation – Connected Vehicles
- Railway Networks
- Power distribution – Smart Grid



**Research collaboration** and dissemination

- Online tools, real-time collaboration
- Versioned models and results
- Model and tool libraries



Created by Angela from Noun Project

VANDERBILT UNIVERSITY

# FOUNDATIONS

**Web-based collaborative interface**

**WebGME – Meta-programmable design environments**

**Integrated simulation tools – CPS domains**

GridLAB-D, SUMO, OMNeT++, veins

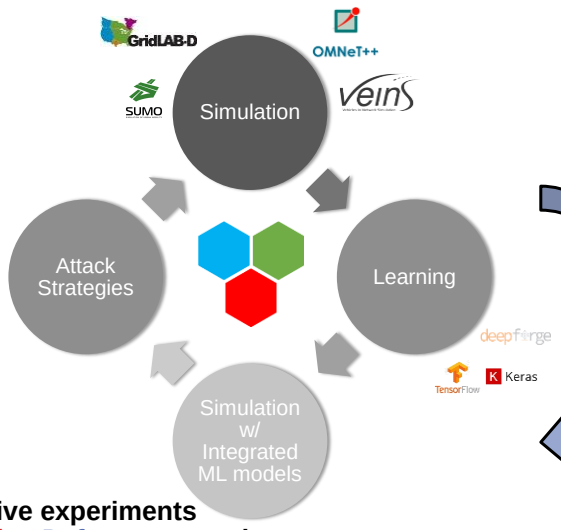
**Machine learning frameworks and tools**

deepforge, TensorFlow, Keras

VANDERBILT UNIVERSITY

3

# DESIGNING RESILIENT CPS WITH INTEGRATED MACHINE LEARNING COMPONENTS




The 4 R's of the life-cycle of ML components

- **R**efresh
- **R**elearn
- **R**evise
- **R**evalidate

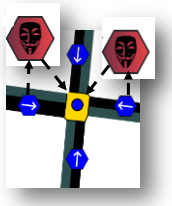
**Iterative experiments**
**Attack – Defense strategies**

VANDERBILT UNIVERSITY

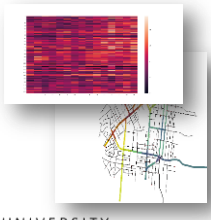
4

# HIGHWAY TRANSPORTATION
## CONNECTED VEHICLES

- **Previous work**
  - Pre-built highway scenarios using SUMO
  - Focus on **highway infrastructure**
  - Attack-library



- **Integrating Veins**
  - Focus on **connected vehicles**
- **Experiment Scenario: Emergency Vehicles**
- **Research Challenge:** Resilient Traffic Forecasting
  - **Queue Estimation**
  - **Flow Prediction**



VANDERBILT UNIVERSITY

5

# RAILWAY NETWORKS

**Multiple levels of abstractions**

- **Low-fidelity traffic simulation SUMO**

- **High-fidelity hardware-in-the-loop (HIL) emulation**

  **BeagleBone Black**

  **CAN bus, ZeroMQ integration**

- **Attack library**

  **Sim: DOS, Delay, Integrity, Corruption**

  **HIL: DDoS**

VANDERBILT UNIVERSITY

6

# POWER DISTRIBUTION SMART GRID

**Smart Grid CPS domain**

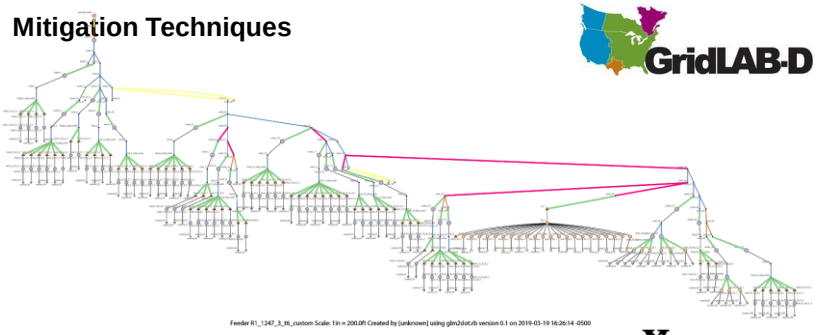
- Distributed power meters (~100 households)

**Load Prediction** with **Deep Neural Networks**

**Adversarial Attack Strategies**

**Mitigation Techniques**

GridLAB-D



VANDERBILT UNIVERSITY

7

# TALK OUTLINE

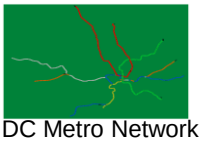
***Three areas of recent research work:***

1. **Railway network simulations with HIL (transitioned to NIST)**
2. **Web-based collaborative modeling and experiments on power distribution grids with market attacks**
   - << DEMO >>
3. **Resilient forecasting of grid loads under stealthy adversarial attacks (uses *DeepForge* for deep-learning)**
   - << DEMO >>

VANDERBILT UNIVERSITY

8

# 1) RAILWAY NETWORK SIMULATIONS WITH HIL



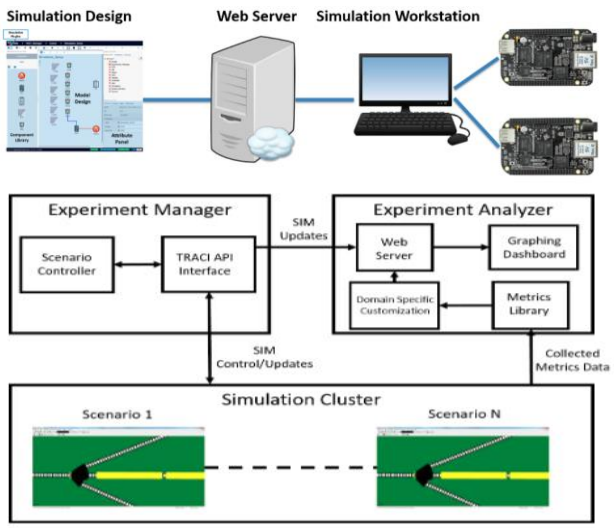
DC Metro Network

- **There have been several safety-critical problems with trains in recent years**
  - Northwest Railway Attack
  - Philadelphia Amtrack
  - Washington State Amtrack
  - South Carolina Amtrack CSX Freight Collision
- **Attackers can leverage interdependencies between physical and cyber domain to affect train behavior**
- **Key challenges:**
  - Autonomous distributed control for optimizing travel times
  - Control algorithms for optimizing railway signal/switch operations
  - Resilient control amidst cyber and/or physical network attacks
- **GOAL: Develop a simulation testbed with HIL that enables enhanced analysis of the resilience of railway networks against cyber and/or physical attacks**

VANDERBILT UNIVERSITY

9

# SIMULATION FRAMEWORK

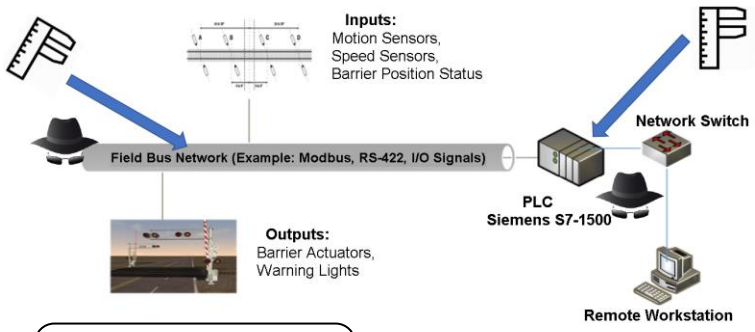



- Support for parallel experiment execution
- HIL support for replacing railway modules with customized controllers
- Real-time results fetching and plotting

VANDERBILT UNIVERSITY

10

# TRANSITION TO NIST




**Network Control Systems Group,**
**Intelligent Systems Division,**
**Engineering Laboratory,**
**NIST Gaithersburg campus**

Neema, Himanshu, Bradley Potteiger, Xenofon Koutsoukos, CheeYee Tang, and Keith Stouffer. "*Metrics-Driven Evaluation of Cybersecurity for Critical Railway Infrastructure*" In 2018 Resilience Week (RWS), pp. 155-161. IEEE, 2018.
Bradley Potteiger, Himanshu Neema, Xenofon Koutsoukos, Chee Yee Tang, Keith Stouffer, "*Simulation Based Rapid Evaluation Platform for Security and Resilience in Railway Infrastructure*" Resilience Week 2019 Symposium, San Antonio, TX, USA, November 4-7, 2019. [submitted]

VANDERBILT UNIVERSITY

11

# 2) POWER DISTRIBUTION GRID SIMULATIONS WITH MARKET ATTACKS

***MODELING CONCERNS***

**POWER-GRID COMPONENTS & DYNAMICS:**
- Generation, Transmission, & Distribution
- Distributed Energy Resources (DERs)
- Distributed Sensors & Controllers
- Multi-Physical Domains & Multi-Rate Dynamics

**ENVIRONMENT & DISASTER RECOVERY:**
- Seasons &Weather
- Recovery & Management for Natural Disasters
- Environment Hazards & Safety
- Special Events w/ Demand Surges

**COMMUNICATION NETWORK & CYBER-ATTACKS:**
- Cyber Network b/n Sensors & Controllers
- Cybersecurity
- Large-scale Data Flows & Real-Time Data Delivery
- Distributed Data Processing

**OPERATIONS & PLANNING:**
- Power Markets & Pricing
- Regulations & Policies (Local & Federal)
- Consumer & Business Models
- Short-Term to Long-Term Planning

***ANALYSIS REQMTS.***

**POWER FLOW ANALYSIS**
- Matching generator-consumer
- Grid stability

**INTEGRATION OF AUTOMATED CONTROL & MARKET FACTORS**
- Optimize utility services
- Reliable & secure operations

**DYNAMIC EFFECTS**
- Demand-Response

**SYSTEM-LEVEL IMPACT ANALYSIS**
- Cyber-attacks
- Resilient cyber-defenses
- Analyzing market approaches for handling renewables

**FOCUS:**

**Transactive Energy Systems**
**Collaborative Modeling & Simulation**
**Resilience against market attacks**

VANDERBILT UNIVERSITY

12

# DESIGN STUDIOS FOR TRANSACTIVE ENERGY SIMULATIONS

**What are Design Studios?**

- Meta-programmable integration platforms
- Domain-specific modeling and experimentation
- Supports collaborative modeling (in *real-time*)
- Web-accessible
- Provide a library of reusable tools and services
- Supports cloud execution of variations of experiments

**1. GridLAB-D Design Studio**

- Pure GridLAB-D simulation (no integration with other simulators)
- Power-grid distribution system simulation + Attacks on markets
- Supports market models for analyzing TE approaches
- Publicly available at: https://cps-vo.org/group/gridlabd

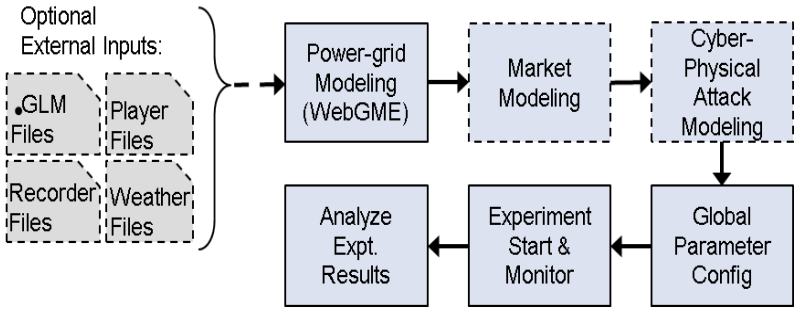
**2. CPSWT-TE Design Studio (CPSWT modified to support TE Co-Simulations)**

- Publicly available at: https://cps-vo.org/group/cpswtte

VANDERBILT UNIVERSITY

13

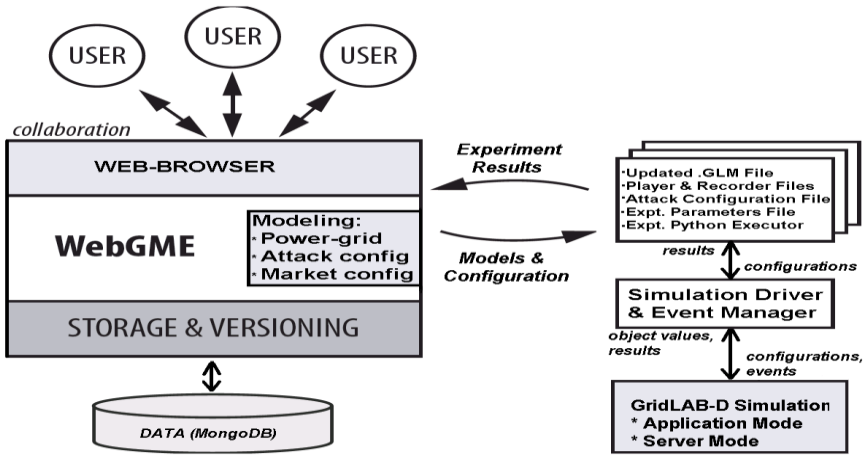
# GRIDLAB-D DESIGN STUDIO: MODELING & SIMULATION WORKFLOW




VANDERBILT UNIVERSITY

14

## GRIDLAB-D DESIGN STUDIO: IMPLEMENTATION ARCHITECTURE






## GRIDLAB-D DESIGN STUDIO: EVENT MANAGER

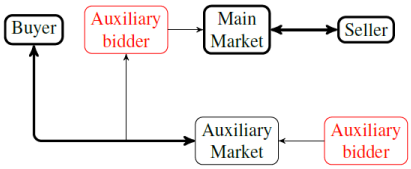
**Algorithm 1** Event manager

**Require:** List of schedules, start time $t_0$, and stop time $t_f$.

1: Pause the simulation at $t_0$.
2: Add future events from schedules in an ordered list.
3: $t \leftarrow t_0$
4: **while** $t \leq t_f$ **do**
5: Wait for a pause in the simulation.
6: Update the simulation time $t$.
7: **for** Each event occurring at time $t$ **do**
8: Execute the event.
9: Update the list of future events.
10: Continue the simulations.

- Between WebGME Driver and GridLAB-D
- Controls GridLAB-D Simulation within a loop
- Reads Object Data
- Updates Object Data

# IMPLEMENTING MARKET ATTACKS IN GRIDLAB-D




**Problems in GridLAB-D:**

1. Bids sent by controllers to the Main Market (MM) cannot be directly modified (can be estimated)
2. Controllers observe the market's clearing prices, and soon as market clearing prices are observed, the reported fake prices get invalidated

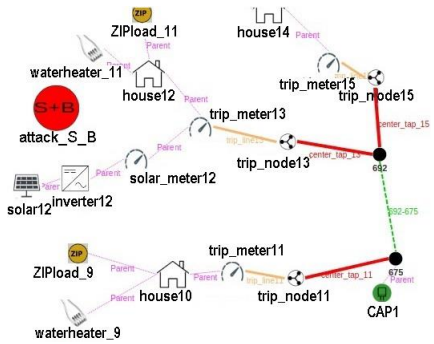
## Our Solution:

- We handle these by introducing Auxiliary Market (AM)
- Buyer bids directly in AM and makes decisions based on the prices in AM
- We estimate the bids because precise bids are not observable
- We implement an Auxiliary Bidder (AB) that sends the estimated bids to MM
- AB also precisely replicates seller's bids in AM (this is possible because in our model seller's bids are constant)

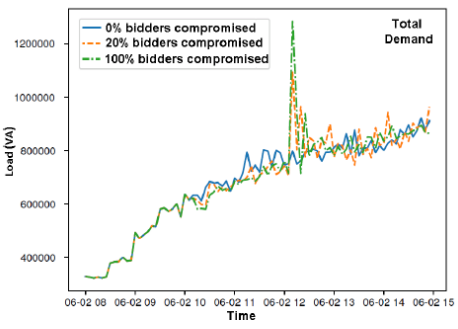
VANDERBILT UNIVERSITY

17

# SAMPLE GRID AND EXPERIMENT




*Part of the distribution grid model*




*Demand-peak created via market attacks*

Himanshu Neema, Harsh Vardhan, Carlos Barreto, Xenofon Koutsoukos, "*Web-Based Platform for Evaluation of Resilient and Transactive Smart-Grids*", 2019 Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES), Montreal, Canada, 04/2019.

VANDERBILT UNIVERSITY

18

# CONCLUSIONS & FUTURE WORK

**CONCLUSIONS**

- Power-grids are a complex system involving many different components
- Increased connectivity and DERs have increased grid's vulnerability to attacks
- A Web-based platform with the following advantages:
  - Web-access
  - Graphical modeling environment
  - Real-time collaboration
  - Transactive energy simulations
  - Experimentation with market attacks
  - Several case study models
  - Library of high-level models

**FUTURE WORK**

- Support design of experiments (DOE)
- High-level model library (more parametric/reusable)
- Experiment results visualization improvements (also plotting live expt. results)
- Support for evaluation of societal-level implications of TE

**<< DEMO >>**



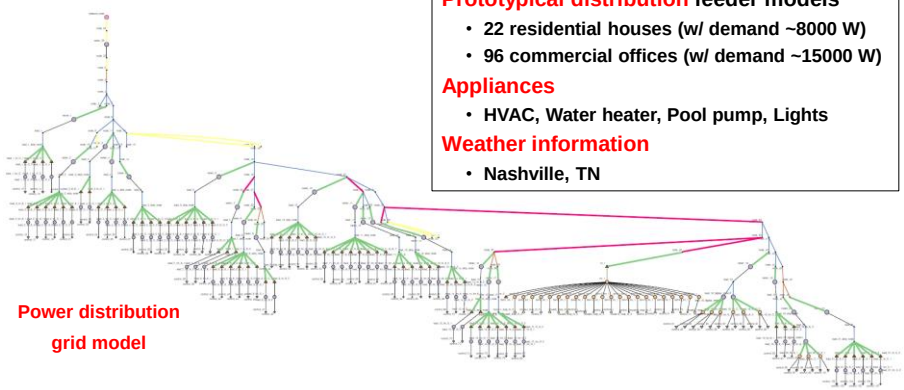
# 3. RESILIENT FORECASTING OF GRID LOADS UNDER STEALTHY ADVERSARIAL ATTACKS

1. **DeepForge for deep-learning**
2. **Rapid design and validation of resilient prediction – defense strategy**

**Prototypical distribution feeder models**
- **22 residential houses (w/ demand ~8000 W)**
- **96 commercial offices (w/ demand ~15000 W)**

**Appliances**
- **HVAC, Water heater, Pool pump, Lights**

**Weather information**
- **Nashville, TN**

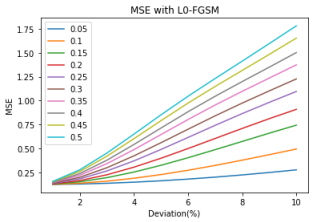


**Power distribution grid model**

# RESILIENT LOAD FORECASTING

Resilient load forecasting for grid operations planning & reliability
(generator schedules, energy reserves, bids in markets, …)

- Uncertainties of weather, user behavior, failures, and renewables
- Attacks on machine learning components




MSE under L0-FGSM attacks.

*Deviation* is the relative amount of perturbation on the sensor readings (attack)

**1. Baseline adversarial attacks**

- $L_0$ and $L_\infty$ constraints: an attacker is allowed to modify a limited number of meter readings each with a max deviation
- The attacker chooses the most sensitive meters using their gradient

**2. Resilient prediction - defense strategy (work in progress)**

- Drop out a different subgroups of meters randomly and observe the variance of the predictions

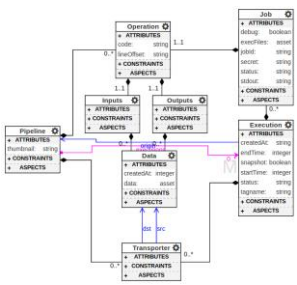
VANDERBILT UNIVERSITY

21

# DEEPFORGE

**Web-based tool for Deep Neural Network research**

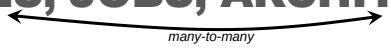
- **Based on WebGME**
- **Generic workflow model**
- **Versioned assets**
- **Keras / Tensorflow**



VANDERBILT UNIVERSITY

22

# DEEPFORGE
## PIPELINES, JOBS, ARCHITECTURES



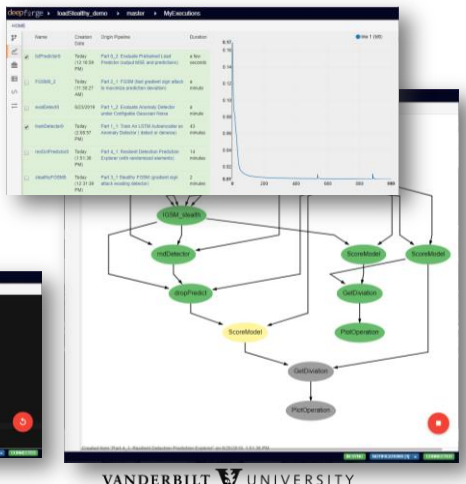

- **Visual workflow model for Python-based operations**
- **Directed Acyclic Graph**
- **Integrated code editor**
- **Input / output / parameter interfaces**
- **References**
  - DNN architecture models
  - Input / output artifacts
- **Executable model**
- **Library of reusable operations**
  - Training
  - **Generic attack models (IGSM, FGSM)**
  - **Generalizable**

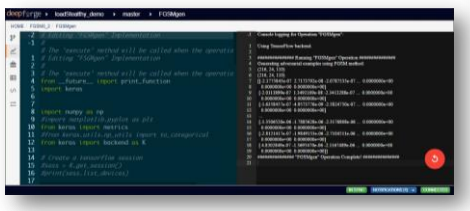

VANDERBILT UNIVERSITY

23

# DEEPFORGE
## PIPELINES AND JOBS

- **Job management and scheduling**
- **Live status updates**
- **Live console output and matplotlib-based graphs**
- **Preserved and versioned results and artifacts**





VANDERBILT UNIVERSITY

24

# DEEPFORGE
## NEURAL NETWORK MODELS

- **Extensible infrastructure: meta-model and code generator**
- **Current support: Torch and Tensorflow / Keras**
- **References from Pipelines (workflow)**

**<< DEMO >>**

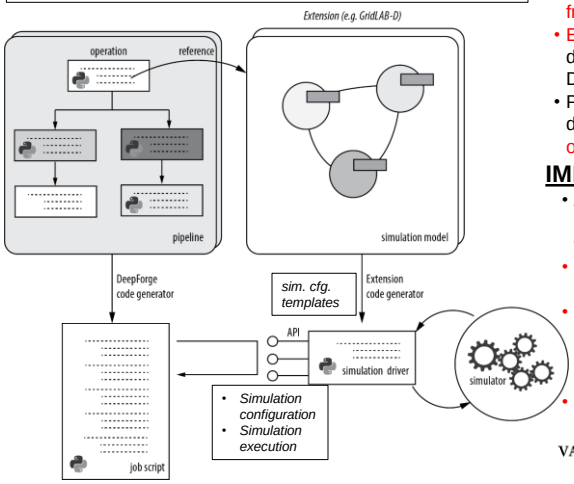
Xingyu Zhou, Yi Li, Carlos Barreto, Jiani Li, Peter Volgyesi, Himanshu Neema, Xenofon Koutsoukos, "*Resilient Forecasting of Grid Loads under Stealthy Adversarial Attacks*". Resilience Week 2019 Symposium, San Antonio, TX, USA, November 4-7, 2019. (submitted)

VANDERBILT UNIVERSITY

25

# INTEGRATED SIMULATION

**FUTURE PLAN:**
**Integrating third party simulators in the workflow**
- Simulation configuration files
- Wrapper code for pipelines




**GOALS**
- Integrate third-party simulators of various CPS domains
- Build on past experience with WebGME based simulation and experimentation front-ends
- Extend pipeline mechanism for integrating domain-specific models and simulators with DeepForge
- Provide an end-to-end environment for designing resilient CPS supporting all 4 R's of the life-cycle of ML components

**IMPACTS**
- Any user-provided python code for simulators can be integrated if it conforms to simulation driver APIs
- Full version-control and sharing of all components for real-time collaboration
- Highly fluid experimentation interface – change simulation model, pipeline embedding, python driver, or simulation configuration
- General-purpose infrastructure with open-source components – reusable/extensible

VANDERBILT UNIVERSITY

26