

# The NebulaStream Platform: Data and Application Management for the Internet of Things

**Steffen Zeuch**, Ankit Chaudhary, Bonaventura Del Monte, Haralampos Gavriilidis, Dimitrios Giouroukis, Philipp M. Grulich, Sebastian Breß, Jonas Traub, Volker Markl



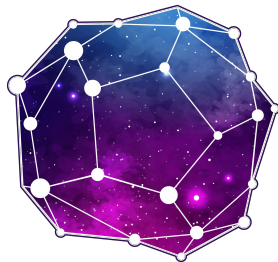
Deutsches  
Forschungszentrum  
für Künstliche  
Intelligenz GmbH



# Our View the IoT:

The Internet of Things (IoT) presents a novel computing architecture for data management:

A geo-distributed, highly dynamic, and heterogeneous environment of massive scale.



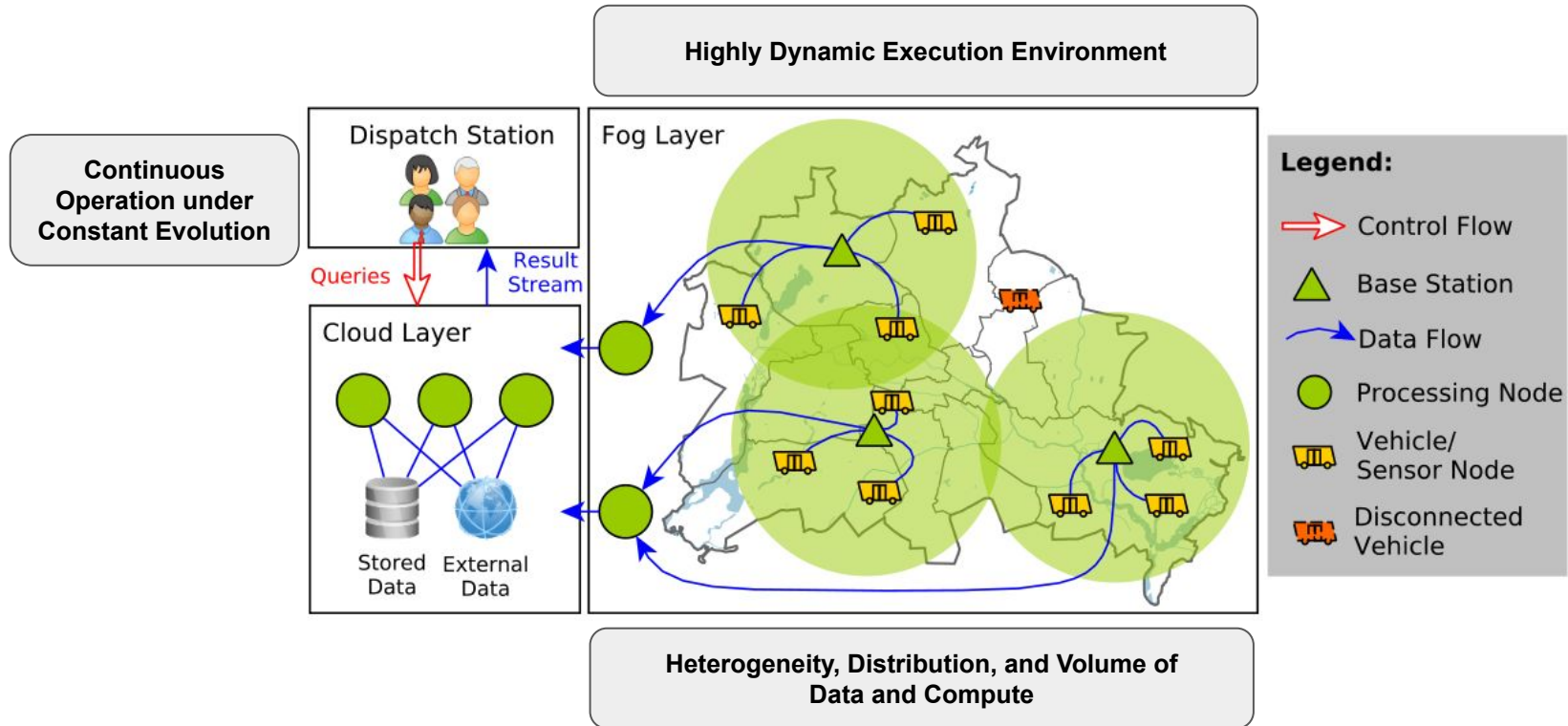
# NebulaStream

An IoT data management system to process thousands of queries over millions of sensors.

# Three questions that everyone should answer before creating a new system

(Is there a need for an IoT system?)

# 1) What's the **new** thing about IoT applications?



## 2) Is there even a **market** for an IoT system?



By 2020, Gartner estimates **internet-connected things will outnumber humans 4-to-1**, creating new dynamics for marketing, sales and customer service.

Leading the IoT  
(Gartner 2017)

The Internet of Things offers a potential economic impact of \$4 trillion to \$11 trillion a year in 2025.

Nine settings  
where value may accrue

Size in 2025, \$ trillion<sup>1</sup>

■ Low estimate ■ High estimate

**Factories**—eg, operations management, predictive maintenance



**Cities**—eg, public safety and health, traffic control, resource management

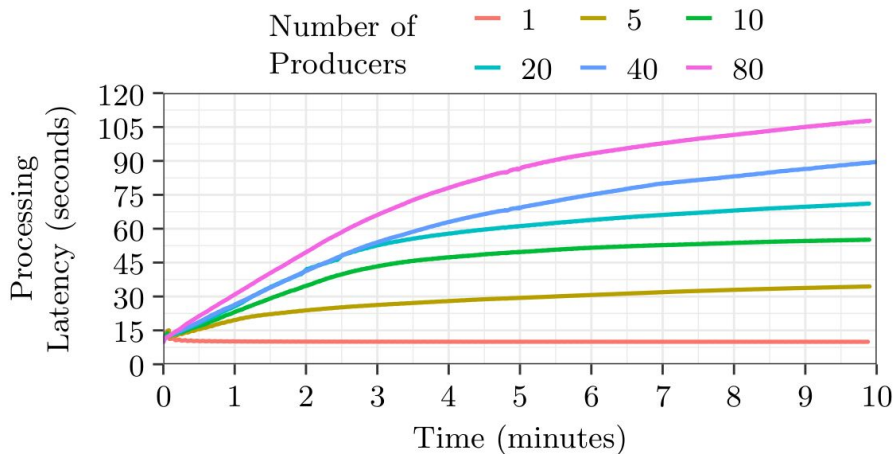
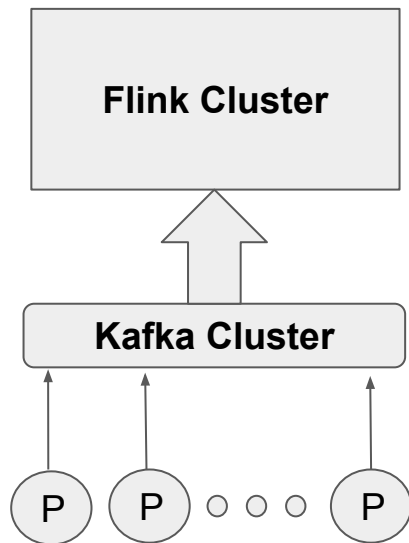


**Human**—eg, monitoring and managing illness, improving wellness



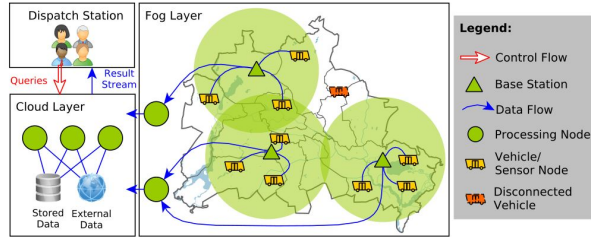
Unlocking the potential of the Internet of Things  
(McKinsey Report 2015)

### 3) Can't we just use **existing** (cloud-)solutions?



**There is no general-purpose, end-to-end data management system for the IoT with similar functionality as Flink or Spark.**

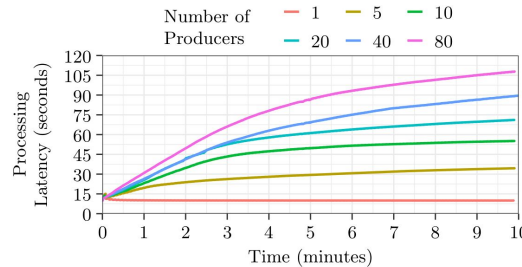
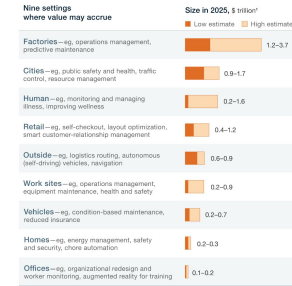
# The answers to the questions are:



Yes IoT is something new

Yes there is a big market/demand

The Internet of Things offers a potential economic impact of \$4 trillion to \$11 trillion a year in 2025.



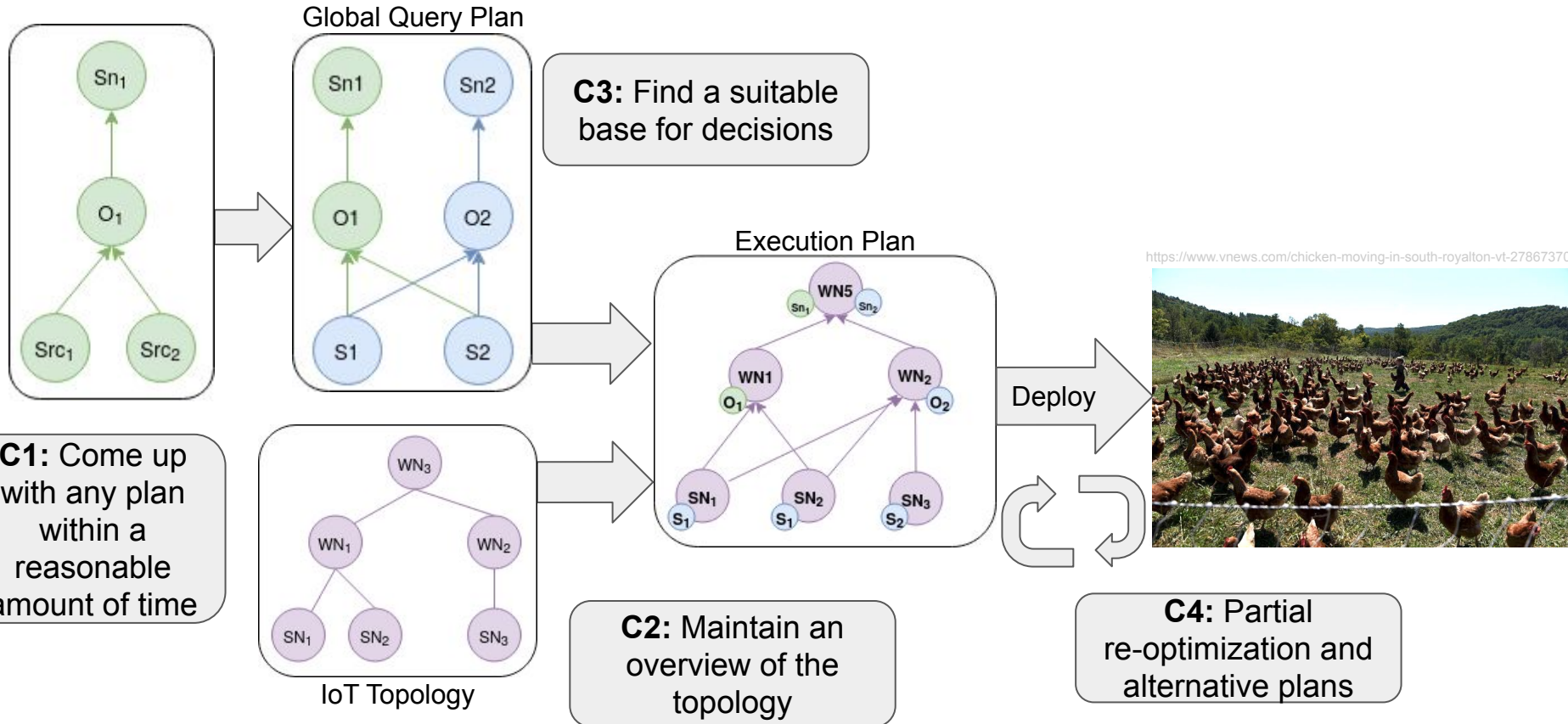
No we cannot just use existing systems



# What are the new research challenges of IoT?



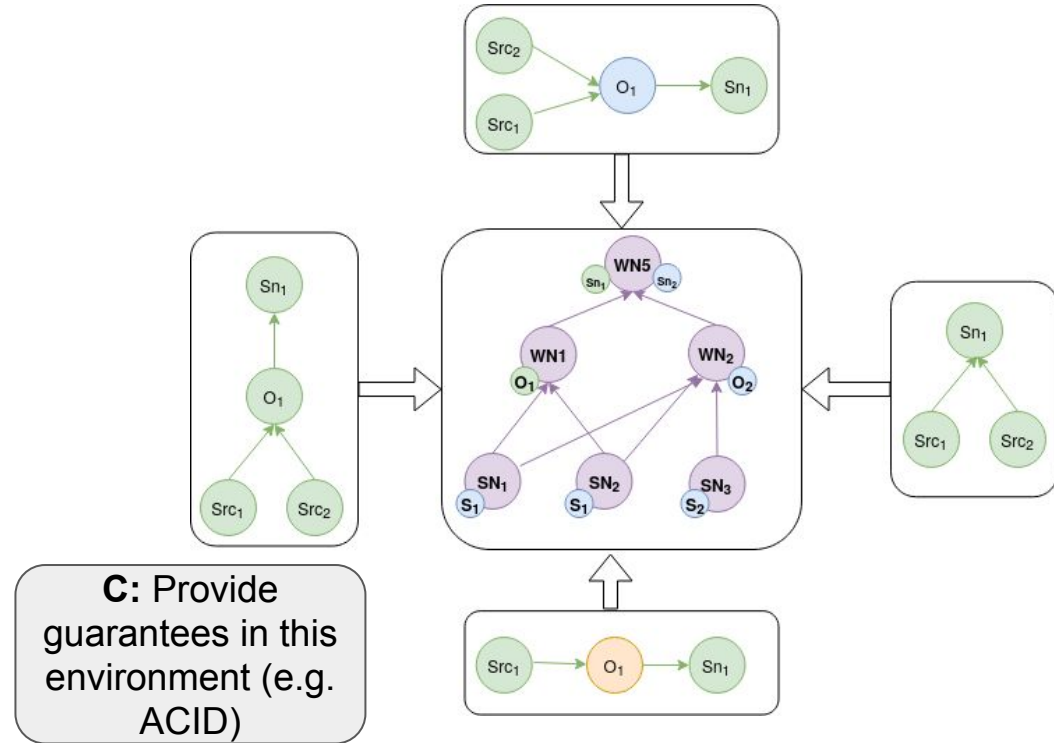
# Query Submission in the IoT



# Do we really have to chase the chickens **centrally**?



<https://www.youtube.com/watch?v=1VFz5Q10vaQ>

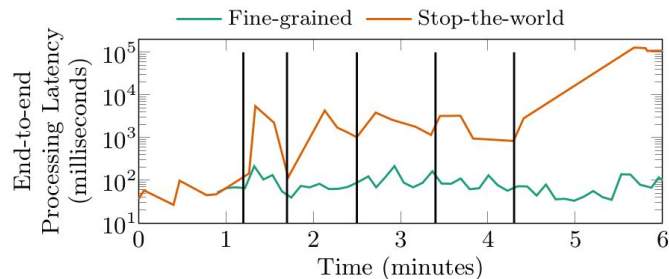


**NebulaStream: A hybrid approach to meet different requirements.**

# Query execution in a **highly dynamic** environment



<https://homediy.eu/schuetzen-sie-ihre-huehner-vor-raubtieren/>



Failures are the common case in IoT.

**NebulaStream: Autonomous node design:**

Local

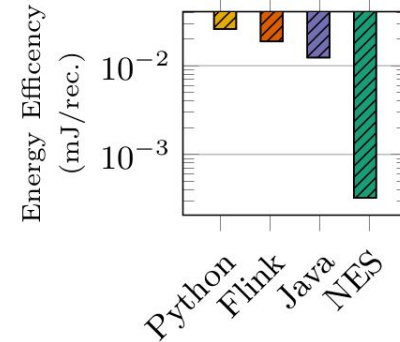
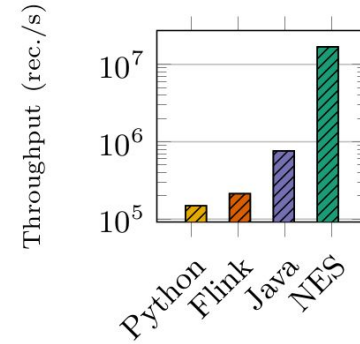
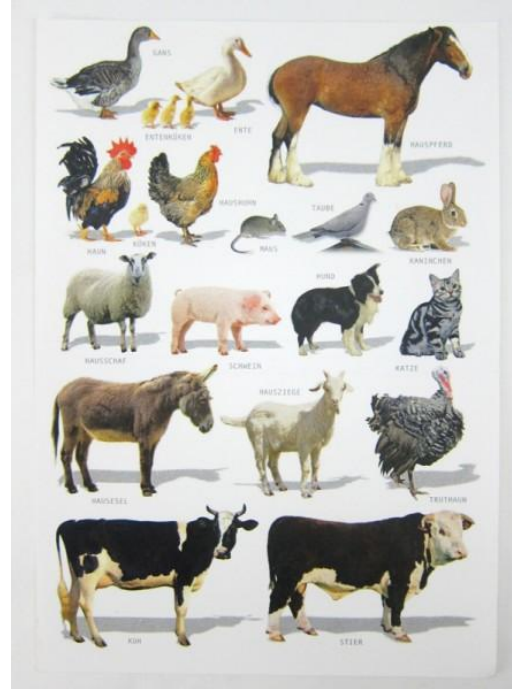
Neighborhood

Global

# How can we cope with a highly heterogeneous IoT?



<https://www.hagemann.de/biologie/zoologie/bestimmungstafel-tiere-auf-dem-bauernhof>



**NebulaStream:** Hardware-tailored code generation and a dynamic execution engine.

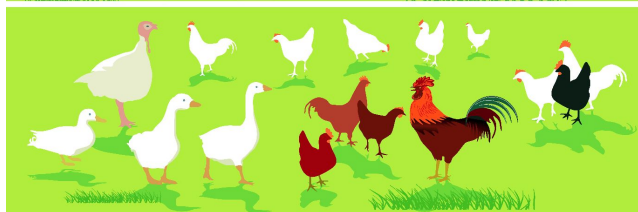
# How to handle “IoT Scale” ?



Cloud Layer



Fog Layer



Sensor Layer

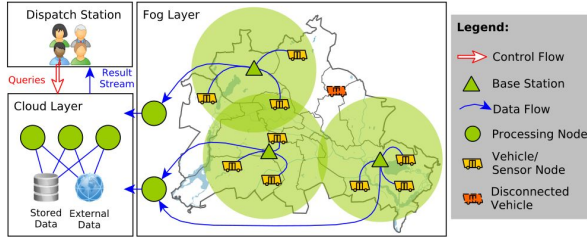


## Key Principles:

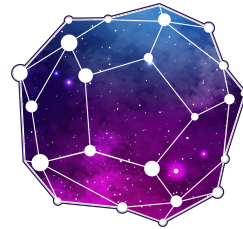
- Reduce complexity
- Reduce data volumes
- Use resources efficiently

**NebulaStream:** A system for a unified sensor-fog-cloud environment.

# Summary



Yes we need a new system



## NebulaStream

A general-purpose,  
end-to-end data  
management system for  
the IoT.

<http://www.nebula.stream>

Let's build the  
next Flink



<https://pixabay.com/de/vectors/tier-vogel-huhn-hen-lustig-159510/>