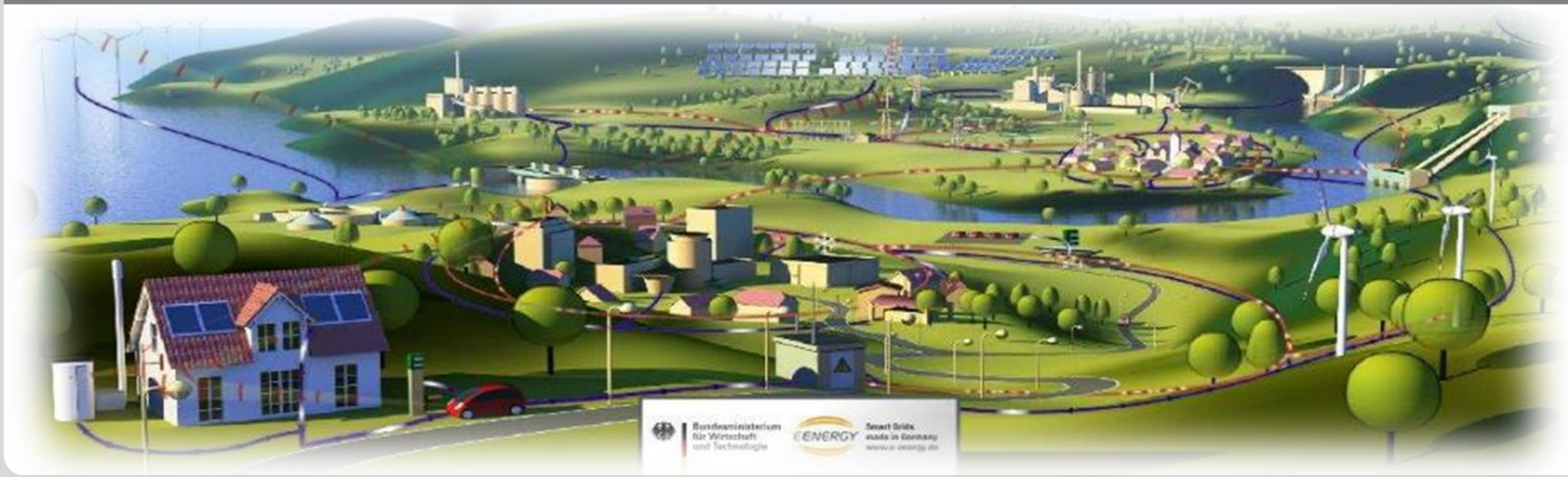
Karlsruhe Institute of Technology

# Concept of a Big Data Infrastructure for Future Power System Control Centers

Eric Braun, Wolfgang Süß, **Karl-Uwe Stucky**, Clemens Düpmeier

INSTITUTE FOR APPLIED COMPUTER SCIENCE (IAI)



# Agenda

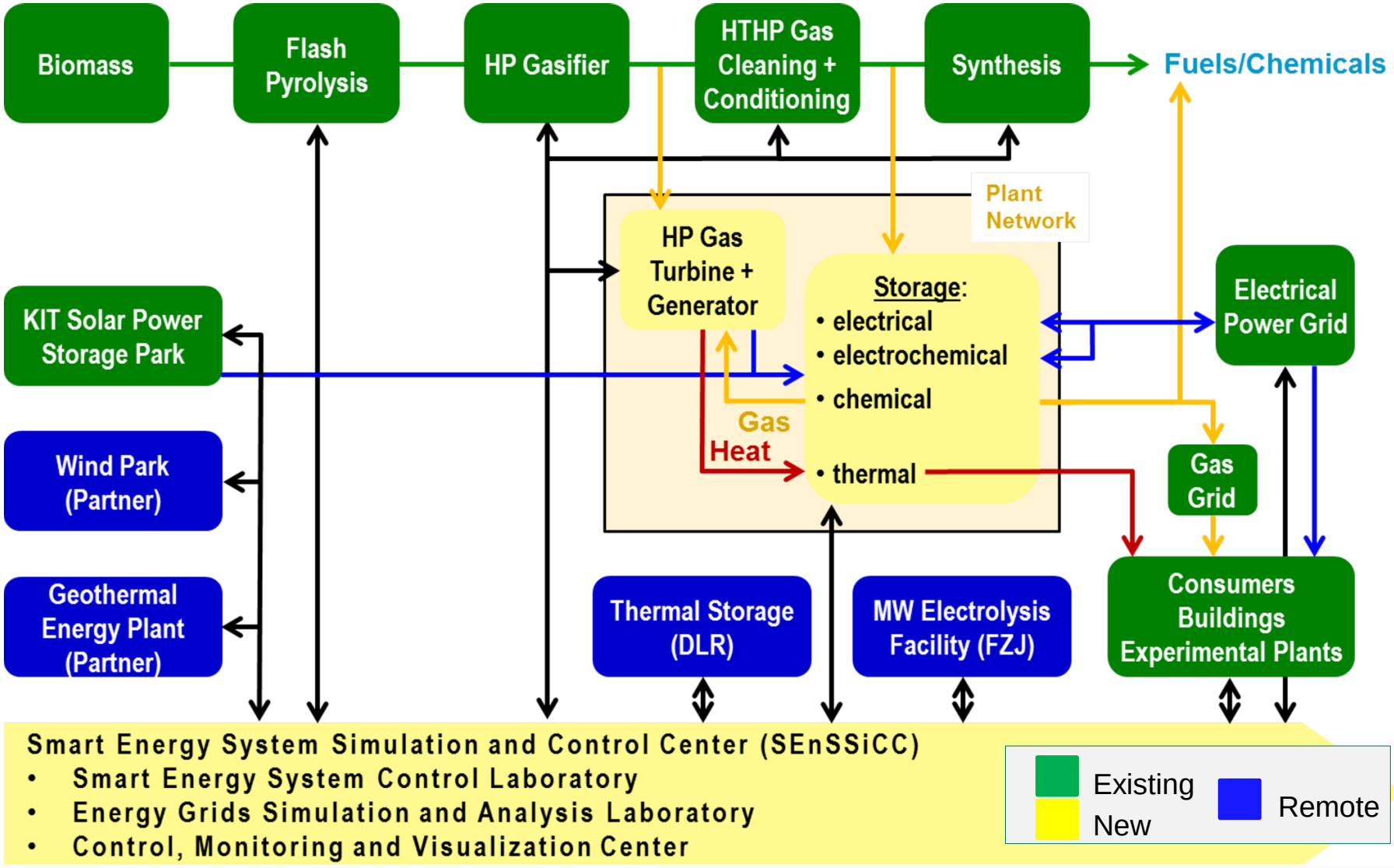
- Goal
- Projects Energy Lab 2.0 and Energy Systems 2050
- Concept Overview
- Details of the Concept

# Goal

- A New Software Architecture for a Distributed, Modular and Scalable IT Infrastructure of Future Power System Control Centers

- Key Challenges:
  - **fast development**
  - distributed cluster-wide **security system** (single sign-on)
  - generic **data management**, **analytics** and **ingestion** (for energy related data)
  - support for **CIM** and **IEC 61850** semantics
  - integration of IEC 61850 **field devices**
  - **high availability**
  - **highly configurable and automated** infrastructure (backend and frontend)

# Energy Lab 2.0 Infrastructure



# Smart Energy System Simulation and Control Center (SEnSSiCC) of Energy Lab 2.0

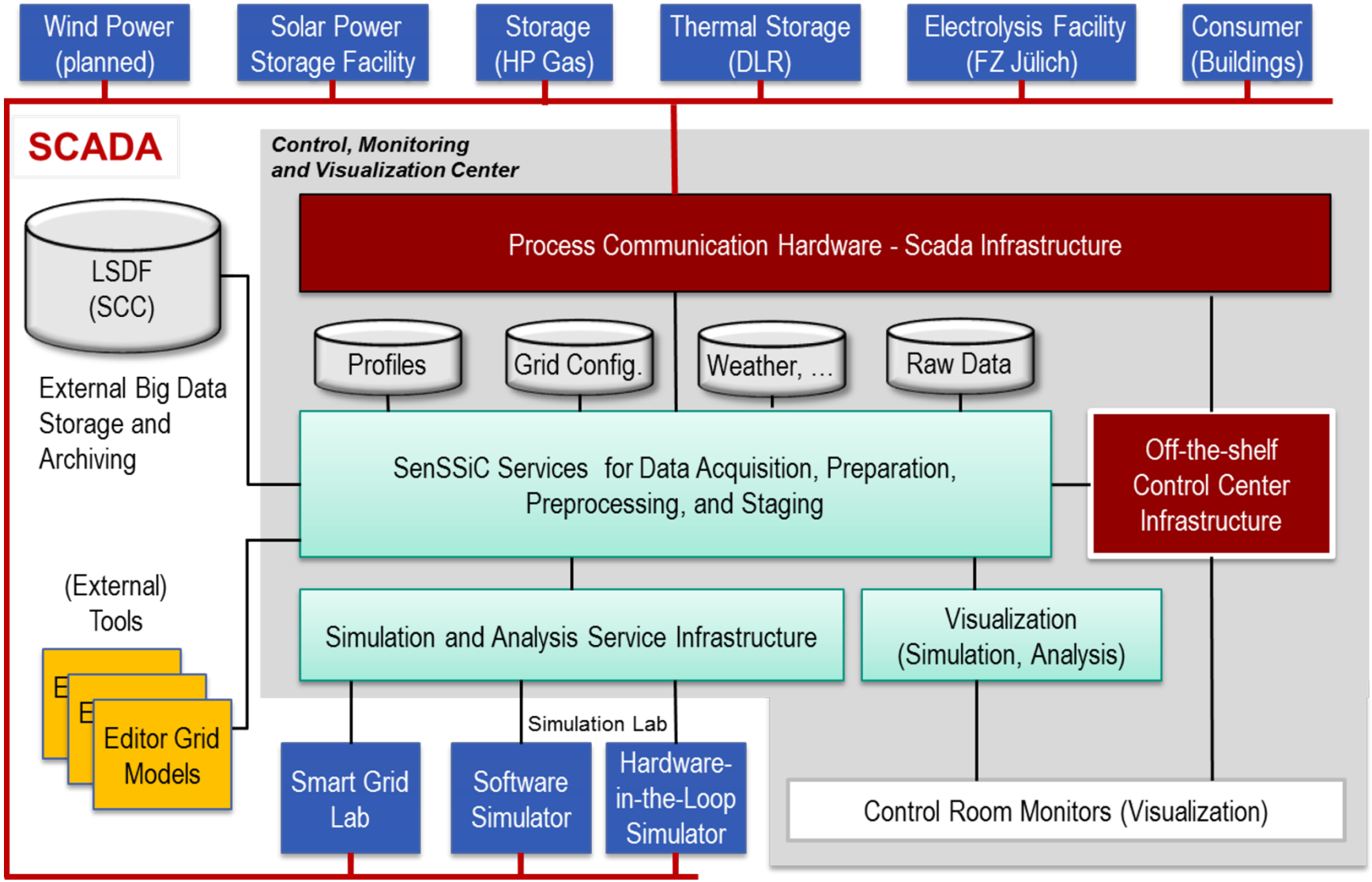
Main parts

- Smart Energy System Control Laboratory with
  - Power hardware in-the-loop test facility
- Energy Grids Simulation and Analysis Laboratory
- Control, Monitoring and Visualization Center (CMVC)

Cross-Cutting Topics

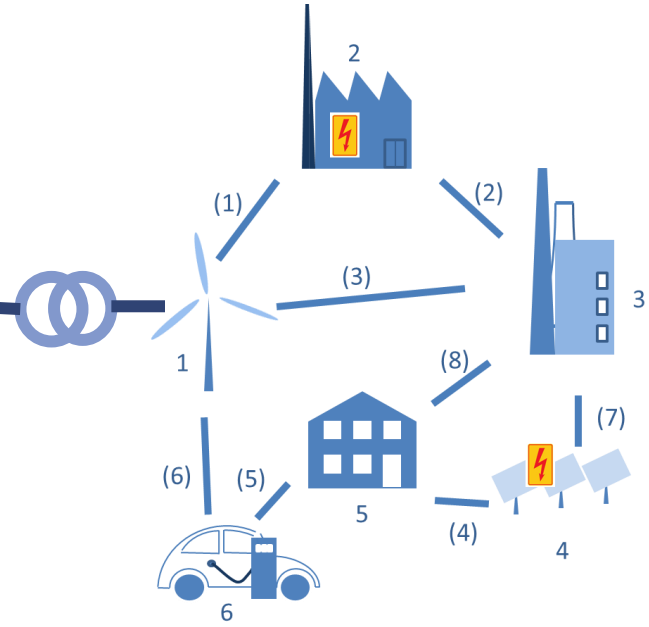
- Big Data
  (Data management
  and analysis)
- Smart control algorithms
- Security, Safety, Controllability

# Control, Monitoring and Visualization Center



# Control, Monitoring and Visualization Center

- Combines own research solutions for monitoring, control and visualization of grid simulations with commercial control center software and a SCADA communication infrastructure
- Integrates grid lab hardware and external Energy Lab plants
- Should look like a real grid control center for operators
- Research on new control center software components and architectures, newer communication technology and risks, tools for demand side management, demand response, grid utility operations



# Energy System 2050

- Running until 2019
- Goals:
  - improving understanding of energy systems
  - developing technological solutions for use by politics and industry.
- Develop toolbox for an „Internet of Energy“ based on Big Data infrastructures
- Use SEnSSiCC infrastructure as test bed

# Energy System 2050

Work in the research topic "Toolbox with Databases" is done under 7 work packages:

- **FT 5.1: Data formats, quality of data and database service**
- **FT 5.2: Building and network models for „Campus Living Labs“**
- **FT 5.3: Simulation platforms, IT architecture and security**
- **FT 5.4: Monitoring and data analysis**
- FT 5.5: Forecasting and automated energy scheduling
- FT 5.6: Optimization and future planning of energy systems
- FT 5.7: Control of future energy systems

# CMVC IT-Infrastructure

Karlsruhe Institute of Technology

**Applications**

Visualization | Energy Management | Outage Management | Simulation of Power Grids

*Frontend Framework*

*Generic Microservice Backend (GMB)*

**Big Data Services**

*Generic Data Services (GDS)*

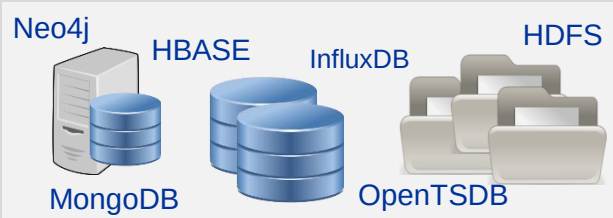
*Process Operation Framework (Proof)*

Data Relation-ship | Schema | Master Data | Time Series Data | Digital Assets

Generic Operations (Data Analysis, Model, Optimization, Simulation)

**Big Data Tools**

***Data Storage Technologies***




***Big Data Analysis Tools***

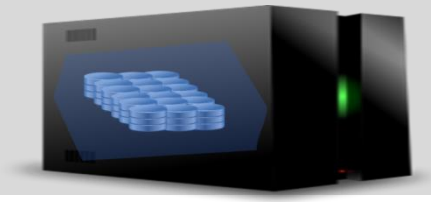
Apache Storm
Apache Spark
Elastic Search
Apache NiFi

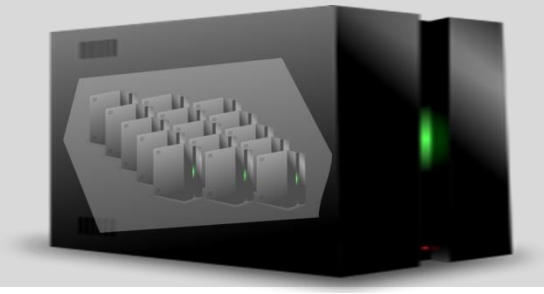
**Cluster Management**

***Distributed Cluster Computing System (DC/OS, MapR, Cloudera, …)***
***Container Virtualization (Kubernetes, Docker, …)***

**Hardware and OS**

***Computing Cluster (Windows and Linux supported)***





# First and Second Layer

Karlsruhe Institute of Technology

**Visualization** | **Energy Management** | **Outage Management** | **Simulation of Power Grids**

***Applications***

***Frontend Framework***

***Generic Microservice Backend (GMB)***

***Generic Data Services (GDS)***

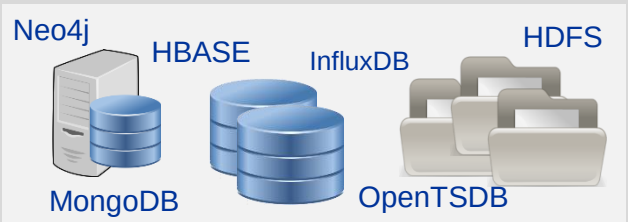
***Process Operation Framework (Proof)***

**Data Relation-ship** | **Schema** | **Master Data** | **Time Series Data** | **Digital Assets**

**Generic Operations (Data Analysis, Model, Optimization, Simulation)**

***Big Data Services***

***Data Storage Technologies***




***Big Data Analysis Tools***

Apache Storm
Apache Spark
Elastic Search
Apache NiFi

***Big Data Tools***

***Distributed Cluster Computing System (DC/OS, MapR, Cloudera, …)***
***Container Virtualization (Kubernetes, Docker, …)***

***Cluster Management***

***Computing Cluster (Windows and Linux supported)***

***Hardware and OS***

# Basic Layers

- First Layer → Hardware and Operating System
  - Computing Cluster
- Second Layer → Cluster Management
  - Big Data Platforms
    - MapR
    - ***Cloudera***
    - DC/OS
  - Includes
    - Virtualization (Docker, Kubernetes)
    - Hadoop File System (HDFS)
    - Apache HBase
    - Yarn
    - Zookeeper

# Third Layer

Karlsruhe Institute of Technology

**Visualization** | **Energy Management** | **Outage Management** | **Simulation of Power Grids** — *Applications*

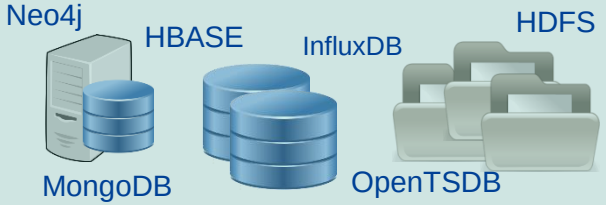
***Frontend Framework***

***Generic Microservice Backend (GMB)***

***Generic Data Services (GDS)*** | ***Process Operation Framework (Proof)***

**Data Relation-ship** | **Schema** | **Master Data** | **Time Series Data** | **Digital Assets** | **Generic Operations (Data Analysis, Model, Optimization, Simulation)** — *Big Data Services*

***Data Storage Technologies***




***Big Data Analysis Tools***

Apache Storm
Apache Spark
Elastic Search
Apache NiFi

*Big Data Tools*

***Distributed Cluster Computing System (DC/OS, MapR, Cloudera, …)***
***Container Virtualization (Kubernetes, Docker, …)***

*Cluster Management*

***Computing Cluster (Windows and Linux supported)***

*Hardware and OS*

# Third Layer → Big Data Tools

- Data Storage
  - Filesystems (HDFS)
  - NoSQL Databases
    - Neo4j
    - MongoDB
    - OpenTSDB, InfluxDB
    - Elasticsearch (*data storage and analysis*)
- Data Analysis
  - Apache Spark
  - Apache Storm
  - Apache NiFi

# Fourth Layer

Karlsruhe Institute of Technology

**Visualization** | **Energy Management** | **Outage Management** | **Simulation of Power Grids** — *Applications*

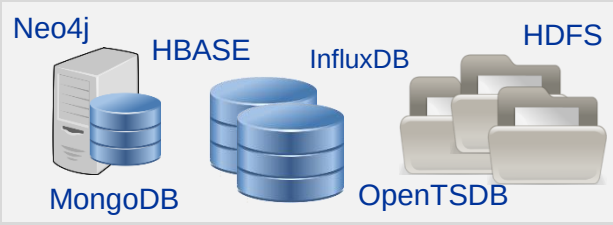
***Frontend Framework***

***Generic Microservice Backend (GMB)***

***Generic Data Services (GDS)*** | ***Process Operation Framework (Proof)***

**Data Relation-ship** | **Schema** | **Master Data** | **Time Series Data** | **Digital Assets** | **Generic Operations (Data Analysis, Model, Optimization, Simulation)** — *Big Data Services*

***Data Storage Technologies***




***Big Data Analysis Tools***

Apache Storm
Apache Spark
Elastic Search
Apache NiFi

*Big Data Tools*

***Distributed Cluster Computing System (DC/OS, MapR, Cloudera, …)***
***Container Virtualization (Kubernetes, Docker, …)***

*Cluster Management*

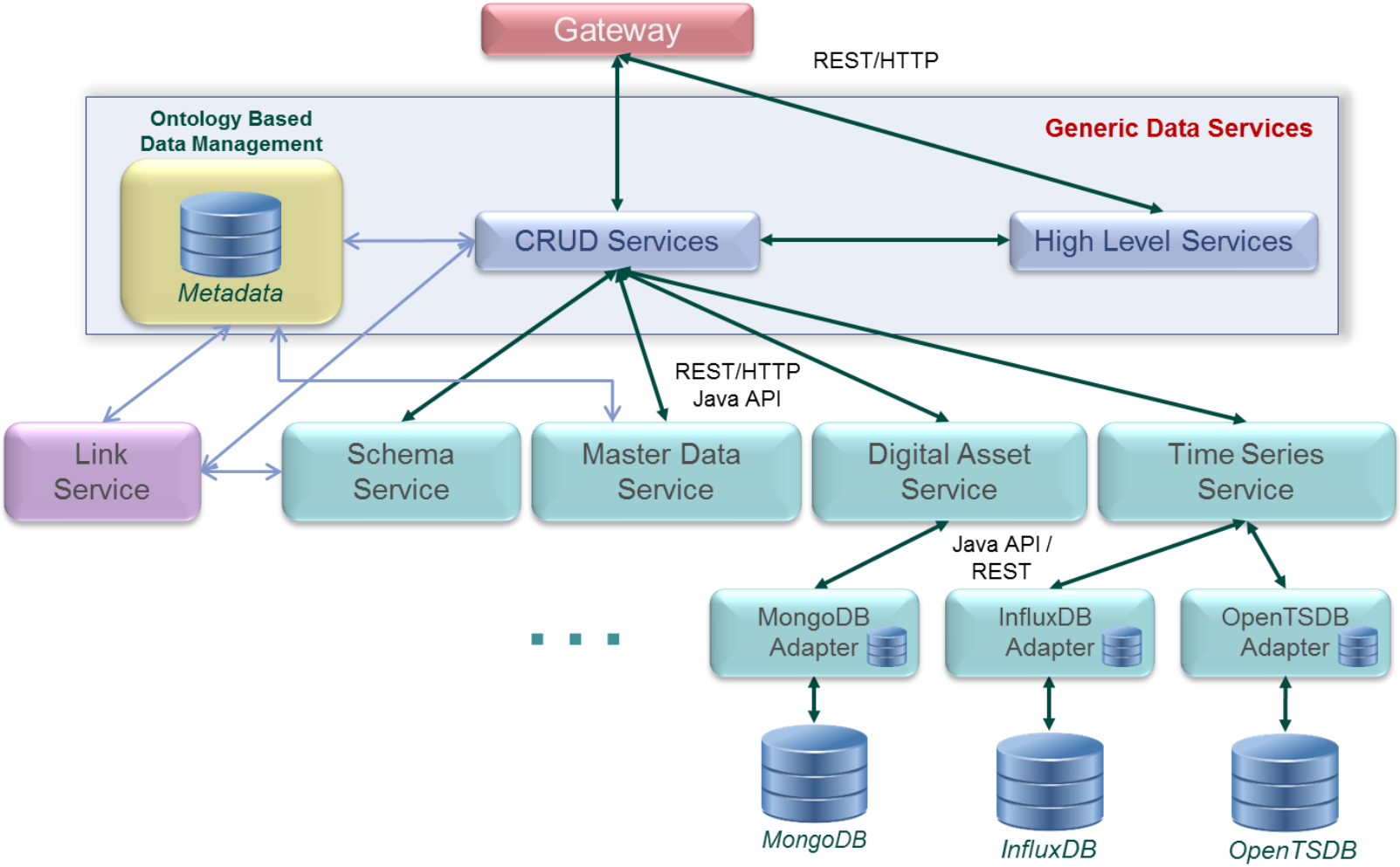
***Computing Cluster (Windows and Linux supported)***

*Hardware and OS*

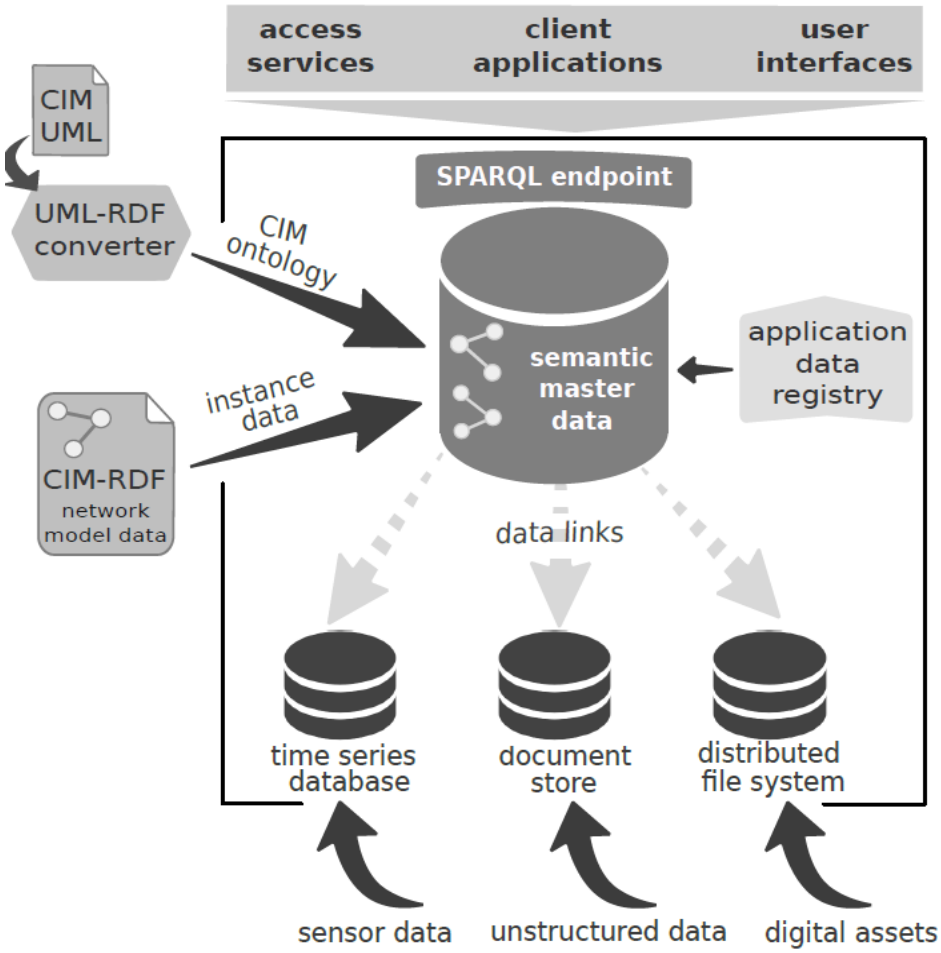
# Fourth Layer → Big Data Services

- Features
  - Service gateway
  - Microservice architecture
- Generic Microservice Backend
  - Security system (single sign-on)
  - Load balancing
  - Service Discovery → Zookeeper
- Components
  - Generic Data Services (GDS) → Data Management
  - Process Operation Framework (Proof) → Data Analysis

# GDS (Generic Data Services)

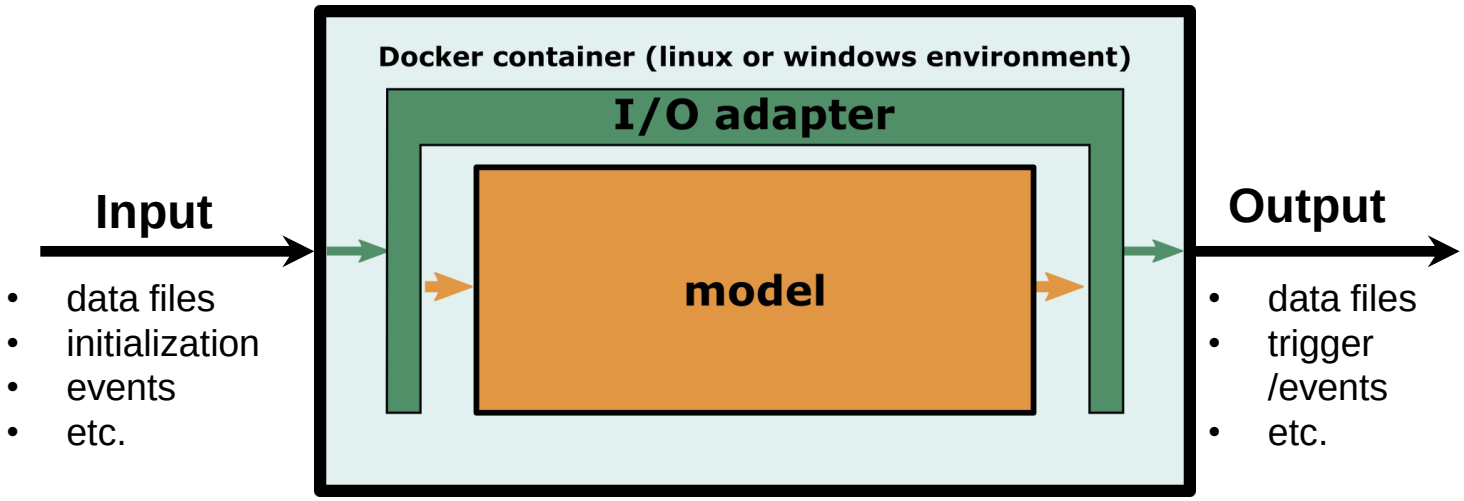


# Common Information Model (CIM)

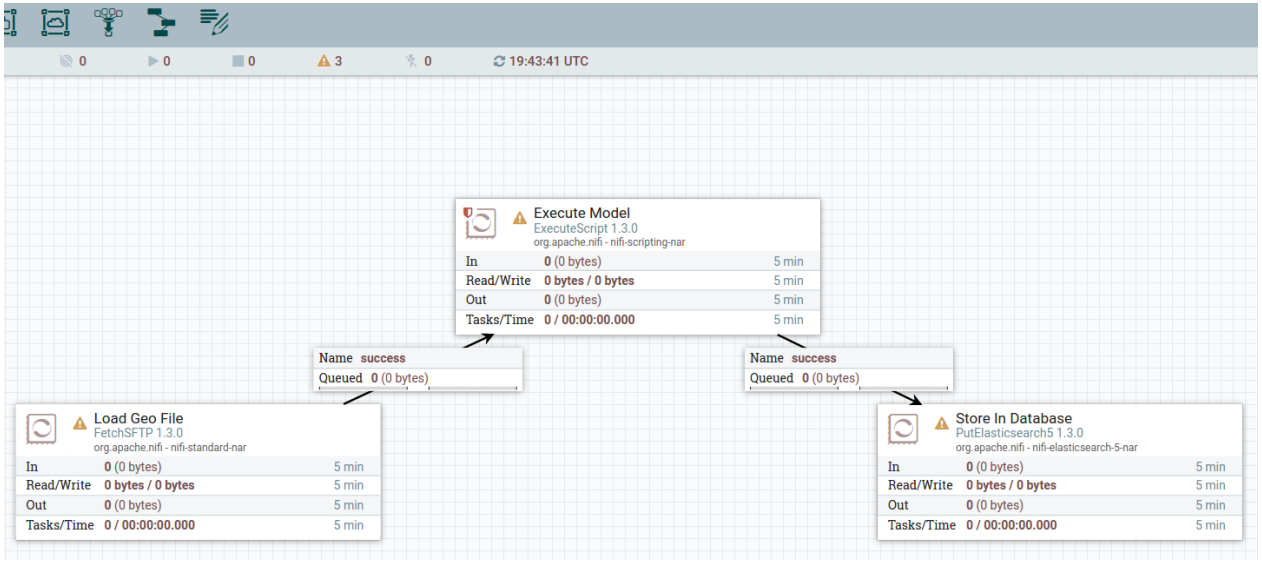



- **Semantic power grid model CIM**
- Data exchange format
- **Ontology-based Data Management (OBDM)**
- Metadata layer

# PROOF (Process Operation Framework)



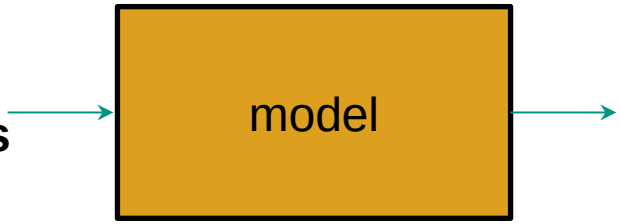

User Interface with Apache NiFi



# PROOF: Example

**Turbines**

**Geo Data**
**Wind Potential Areas**




# Fourth Layer

Karlsruhe Institute of Technology

**Visualization** | **Energy Management** | **Outage Management** | **Simulation of Power Grids** | ***Applications***

***Frontend Framework***

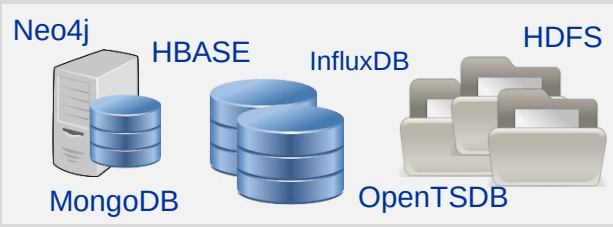
***Generic Microservice Backend (GMB)***

***Generic Data Services (GDS)*** | ***Process Operation Framework (Proof)***

**Data Relation-ship** | **Schema** | **Master Data** | **Time Series Data** | **Digital Assets** | **Generic Operations (Data Analysis, Model, Optimization, Simulation)**

***Big Data Services***

***Data Storage Technologies***




***Big Data Analysis Tools***

Apache Storm
Apache Spark
Elastic Search
Apache NiFi

***Big Data Tools***

***Distributed Cluster Computing System (DC/OS, MapR, Cloudera, …)***
***Container Virtualization (Kubernetes, Docker, …)***

***Cluster Management***

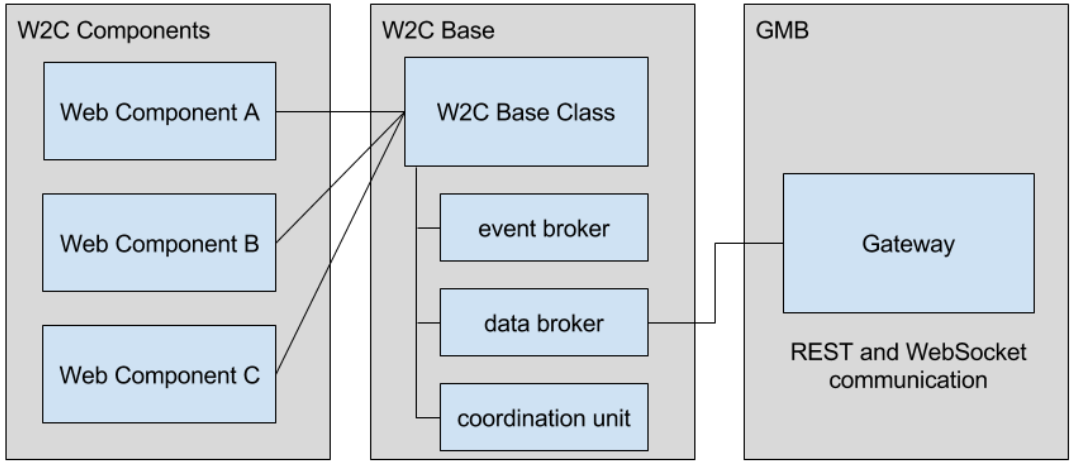
***Computing Cluster (Windows and Linux supported)***

***Hardware and OS***

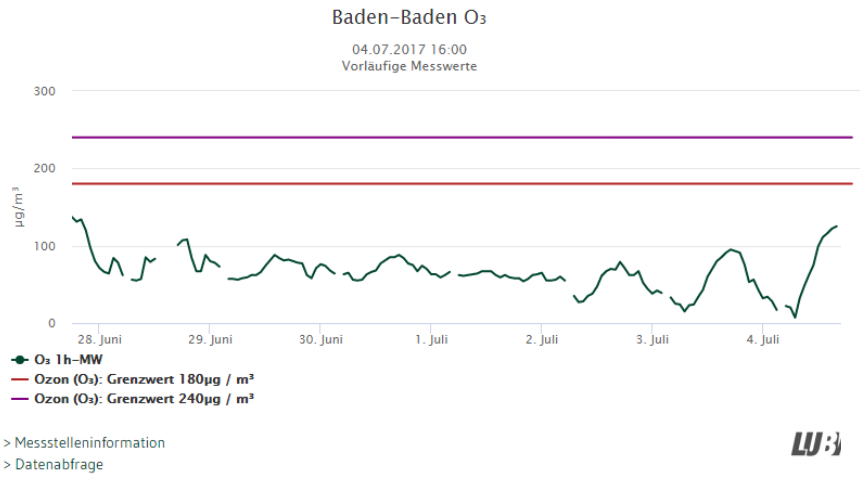
# Lightweight Web Component Framework

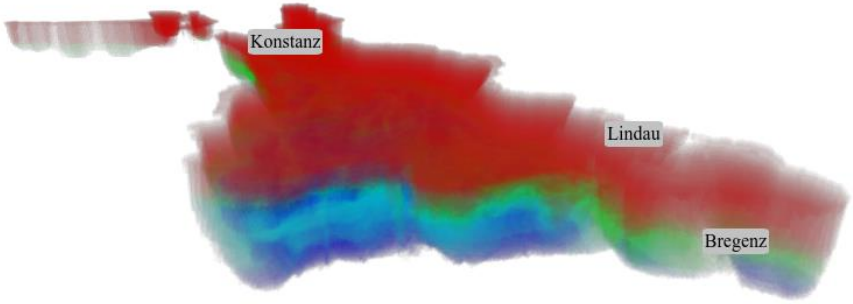
- based on web components technology
  - HTML 5
  - Polymer

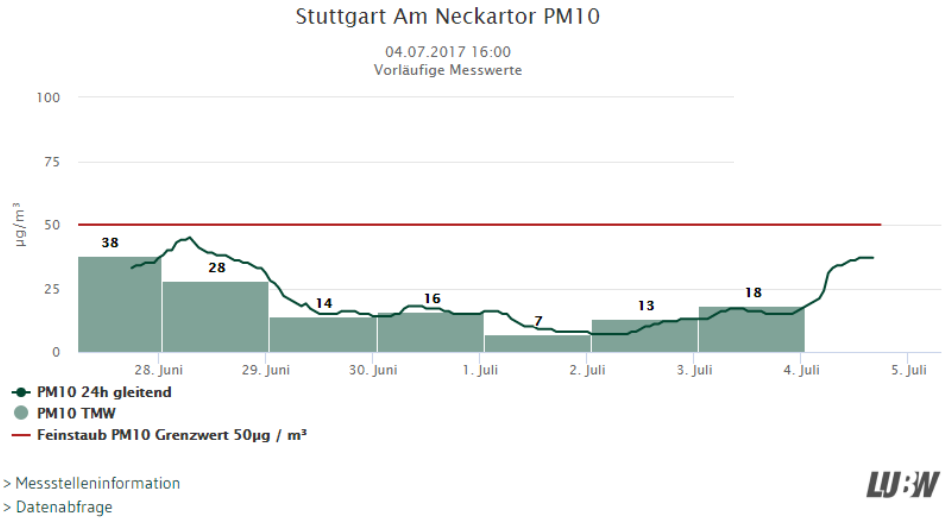

- concept
  - base classes for event and data communication



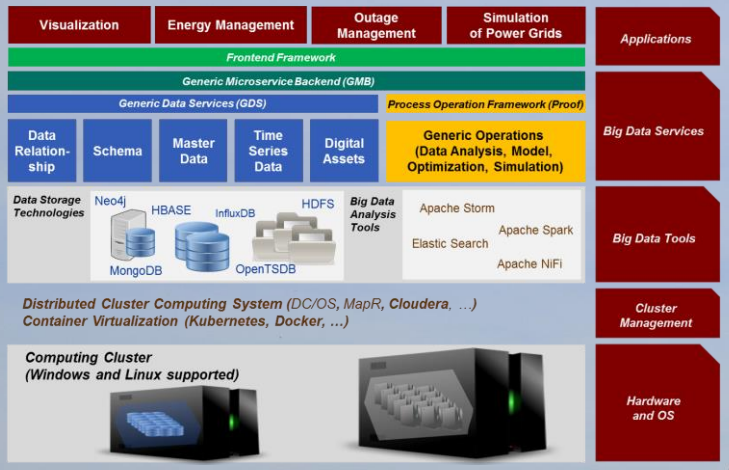
# Example: Visualization



# Thank you!


Visualization
Energy Management
Outage Management
Simulation of Power Grids
Applications
Frontend Framework
Generic Microservice Backend (GMB)
Generic Data Services (GDS)
Process Operation Framework (Proof)
Data Relation-ship
Schema
Master Data
Time Series Data
Digital Assets
Generic Operations (Data Analysis, Model, Optimization, Simulation)
Big Data Services
Data Storage Technologies
Neo4j
HBASE
InfluxDB
HDFS
MongoDB
OpenTSDB
Big Data Analysis Tools
Apache Storm
Apache Spark
Elastic Search
Apache NiFi
Big Data Tools
Distributed Cluster Computing System (DC/OS, MapR, Cloudera, ...)
Container Virtualization (Kubernetes, Docker, ...)
Cluster Management
Computing Cluster
(Windows and Linux supported)
Hardware and OS