Methods of Cloud Computing

Chapter 1: Introduction

Complex and Distributed Systems
Faculty IV
Technische Universität Berlin

Operating Systems and Middleware
Hasso-Plattner-Institut
Universität Potsdam

Who are we? (1/2)

- Research group “Complex and Distributed IT Systems”
  - Lead by Prof. Dr. Odej Kao
  - TEL building at Ernst-Reuter-Platz
  - Secretary: Jana Bechstein, TEL 12-08
  - http://www.cit.tu-berlin.de

Who are we? (2/2)

- Dr. Lauritz Thamsen
  - Postdoc at CIT
  - Responsible for lecture
  - lauritz.thamsen@tu-berlin.de

- Anton Gulenko
  - PhD student at CIT
  - Responsible for exercises
  - anton.gulenko@tu-berlin.de

Structure of the Course

	Time	Room
Lecture	Tuesday 12:00h	H 2013
Exercises	Monday 14:00h	MA 004

- Lecture
  - Every Tuesday
- Exercises
  - Mondays, as needed
  - First of four exercises: 12.11. (further dates will be announced on ISIS)

Examination

- Portfolio exam
  - 60% written exam at the end
  - 40% for the four exercises, each graded with 10%
  - Following “Notenschlüssel 1”
- Written exam: 20.2.2019, 9:00-11:00h
- Exercises will cover concrete usage of IaaS and PaaS technologies and services

Lecture Resources

- ISIS course for announcements, slides, discussions, and exercise groups: https://isis.tu-berlin.de/course/view.php?id=14634

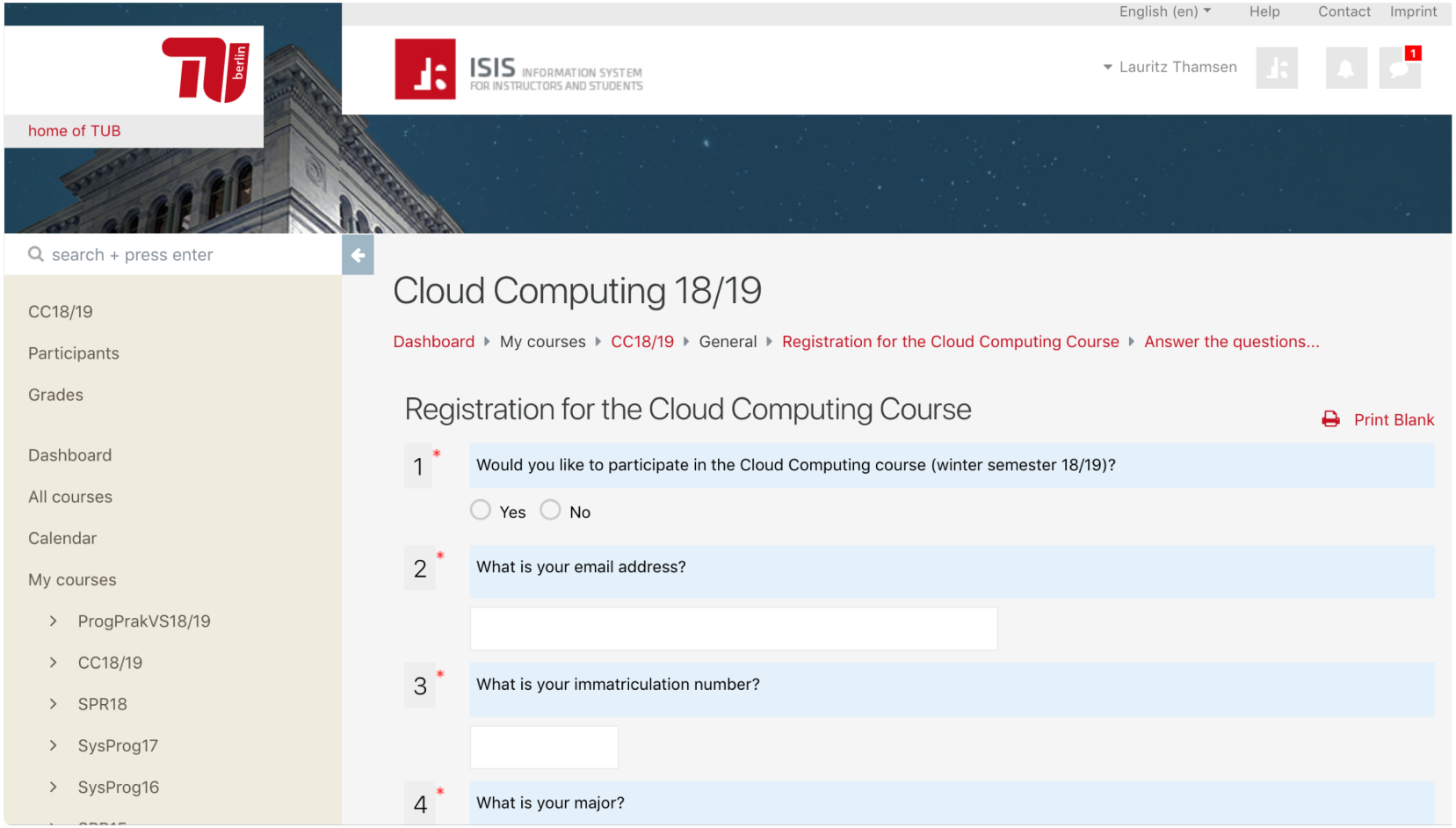
- Please help yourself and each others!

- Recommended reading:
  - Tanenbaum: Modern Operating Systems
  - Humble, Farley: Continuous Delivery
  - Kleppmann: Designing Data-Intensive Applications
  - ... more will be announced during the course!

Module “Cloud Computing”

- Credits (according to ECTS): 6 points
- Limited to 120 participants
- Registration
  - Sign up for the questionnaire on ISIS until Sunday 23.59h
  - We will then follow the official procedure for selecting participants and will notify you on Monday
  - Notified participants can then register for the exam on QISPOS

Registration on ISIS



The Audience

- Distributed Systems?
- Virtualization?
- Fault-Tolerance?
- Big Data frameworks?
- Scale-up vs. scale-out?

Background

Clusters

- Many computers in one room

Grids

- Loosely coupled computers or clusters all over the world

Clouds

- IT resources as a utility

Edge

- Extending the Cloud to the edge of the network

Clusters

- Mostly homogeneous compute resources and software stacks
- Interconnected by a low-latency and high-bandwidth network
- Goal: improving availability, price/performance, resource utilization

- Examples
  - Analytics cluster at Facebook

Grids

- Heterogeneous compute resources
- Connected via a slow network (i.e. the internet)
- Heterogeneous software stacks unified by a middleware

- Examples
  - Worldwide LHC Computing Grid
  - BOINC

NIST Definition of Cloud Computing

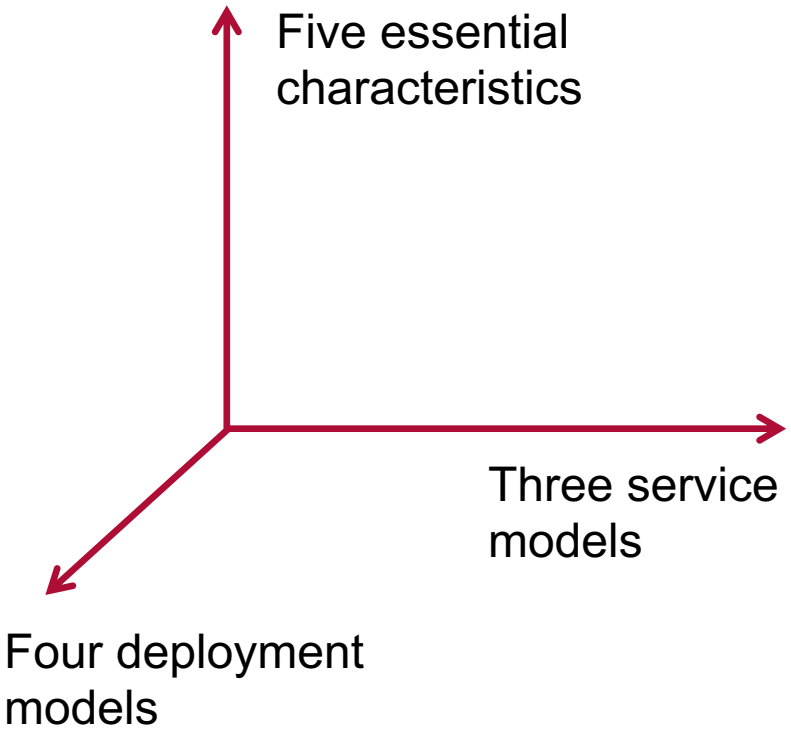
- NIST: National Institute of Standards and Technology
  - Agency of U.S. Department of Commerce
  - Responsible for standardization processes

- Definition of Cloud Computing according to NIST [1]:

> "Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.
>
> This cloud model is composed of **five essential characteristics**, **three service models**, and **four deployment models**."

Dimensions of Cloud Computing (NIST)

- NIST: “Cloud model is composed of
  - Five essential characteristics
  - Three service models
  - Four deployment models”



Five Characteristics of Cloud Computing (NIST)

- On-demand self-service
  - No human interaction required for resource provisioning
- Broad network access
  - Accessible over network with standard mechanisms
- Resource pooling
  - Pooled resources dynamically shared among several consumers, location independence
- Rapid elasticity
  - Capabilities can be provisioned/released on demand
- Measured service
  - Resource usage is monitored, controlled, and reported

Three Service Models of Cloud Computing (NIST)

- Software as a Service (SaaS)
  - Provider‘s application runs on cloud infrastructure
  - Consumer can access application over the network
  - Consumer does not control/manage underlying infrastructure
- Platform as a Service (PaaS)
  - Consumer can deploy custom application onto cloud infrastructure using programming languages, libraries, services and tools supported by provider
  - Consumer does not control/manage underlying infrastructure
- Infrastructure as a Service (IaaS)
  - Provider provisions processing, storage, network resources to consumer
  - Consumer does not control/manage underlying infrastructure but has control over operating systems, storage, and deployed applications

Four Deployment Models of Cloud Computing (NIST)

	Private Cloud	Community Cloud	Public Cloud	Hybrid Cloud
User of the cloud infrastructure?	Single organization	Organizations with shared concerns	Open for the general public	Composition of private/community/public cloud: • Remain distinct entities • Bound together by standard mechanisms • Goal: Enable data / application portability
Owner of the cloud infrastructure?	Organization, third party, combination thereof	Organizations, third party, combination thereof	Business, academic, government organization, combination thereof	
Location of the cloud infrastructure?	On premise, off premise	On premise, off premise	On premise of cloud provider	

Examples

- AWS S3
- AWS EC2
- DigitalOcean
- Google Kubernetes Engine
- Azure Stream Analytics
- AWS Lambda
- AWS SageMaker
- IBM Watson

- Google Mail, Google Docs, MS Office 365…

Background

Clusters

- Many computers in one room

Grids

- Loosely coupled computers or clusters all over the world

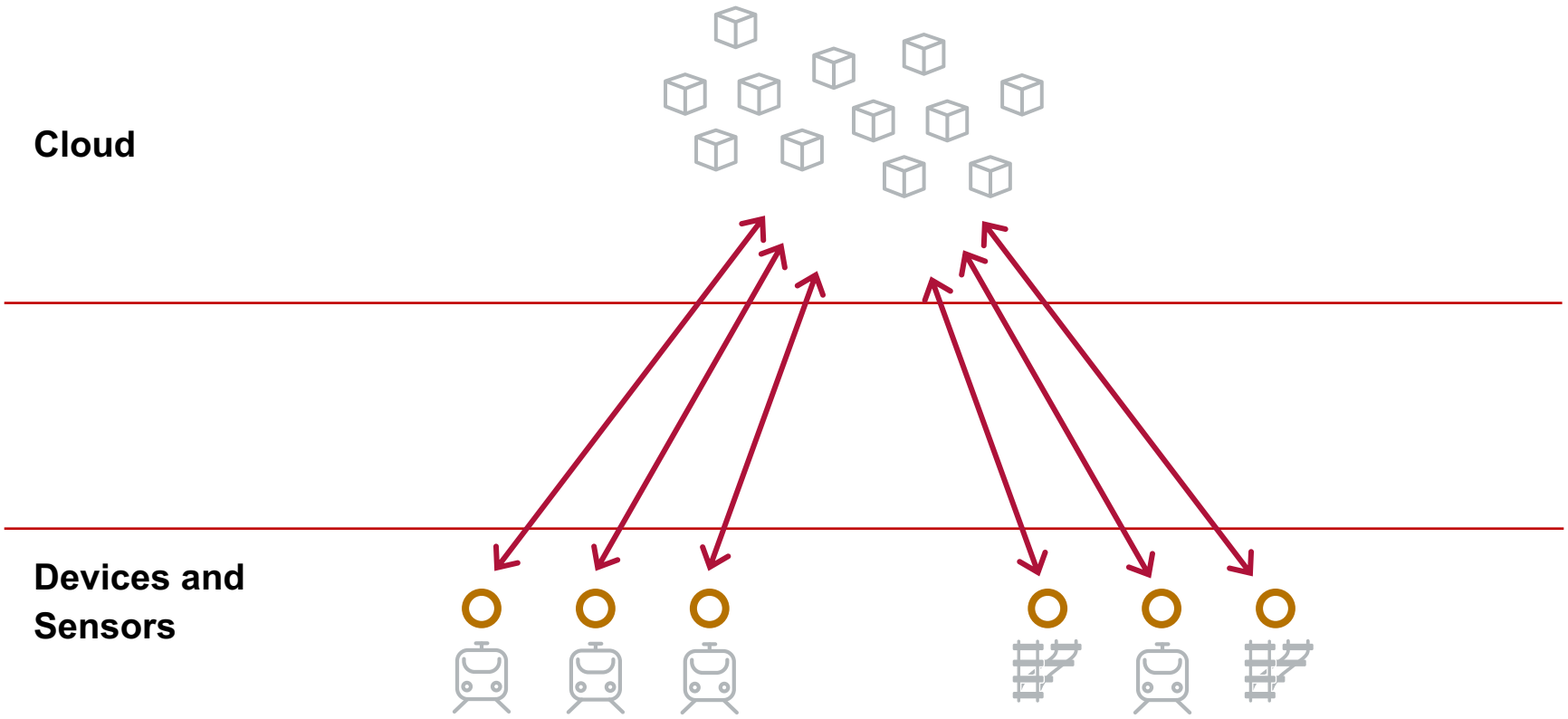
Clouds

- IT resources as a utility

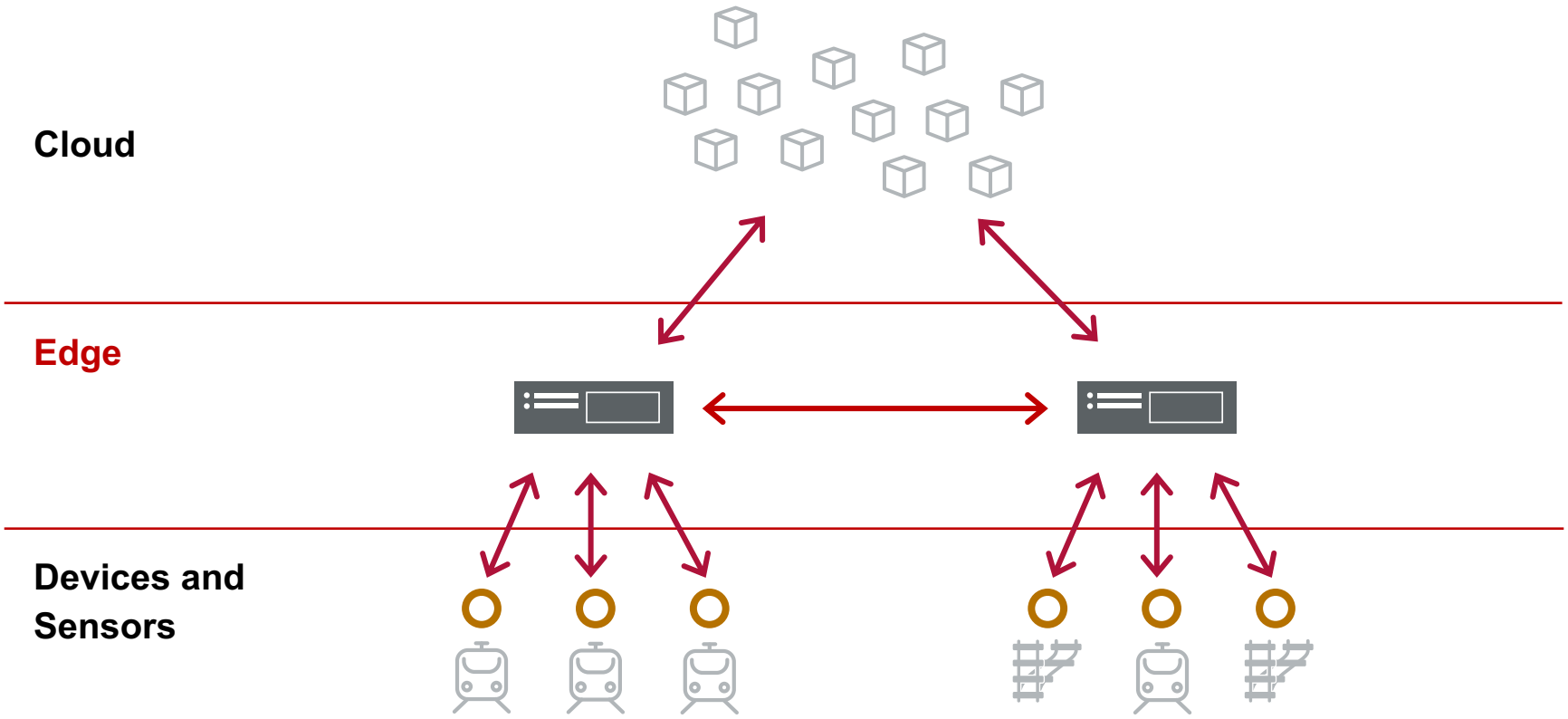
Edge

- Extending the Cloud to the edge of the network

Edge Computing



Edge Computing

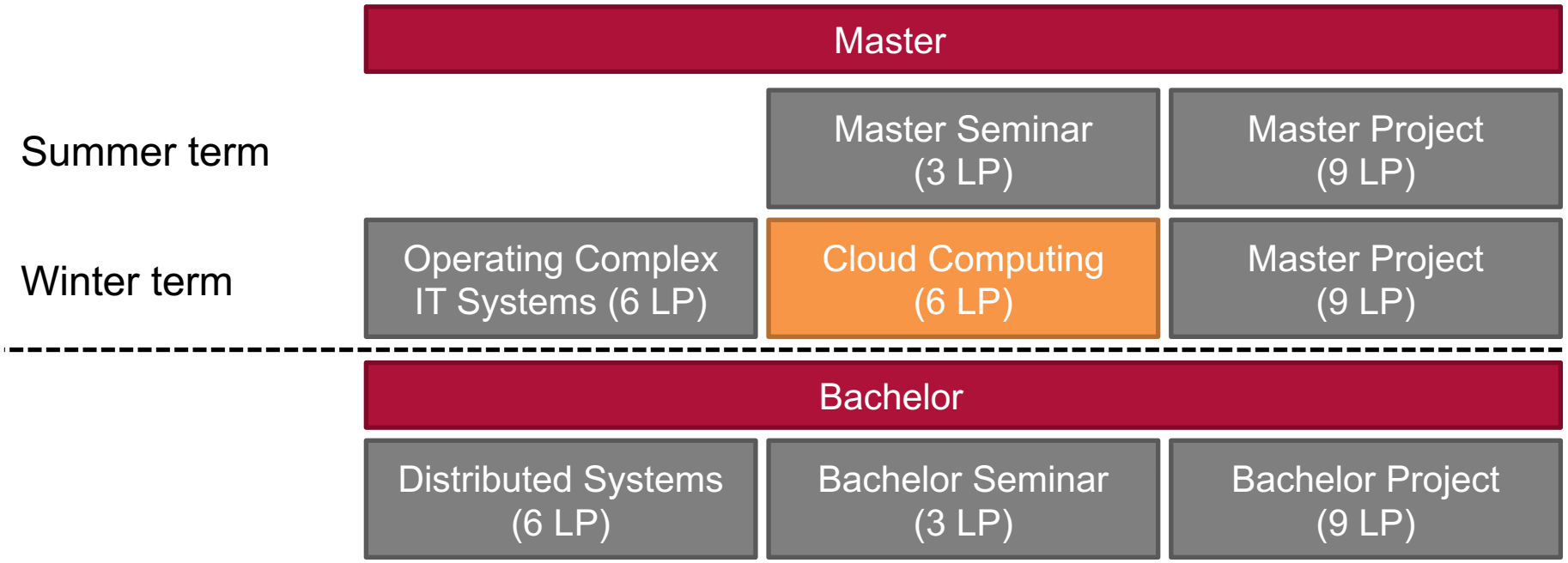


What the Lecture will Discuss

- Goal: Understand what Cloud Computing entails on a technical level
  - Understanding the different levels of abstractions
  - Understanding the implications of resource sharing
  - Learning to take advantage of Cloud platforms
- Content
  2. Virtual Resources
  3. Managing Cloud Resources
  4. Scalable and Fault-Tolerant Applications
  5. Data-Intensive Applications
  6. Platforms and Serverless Computing
  7. Federations, Edge, IoT

Teaching at CIT

- CIT curriculum
  - Bachelor: Basics of distributed systems
  - Master: Advanced topics and stronger research focus



Research at CIT

- Three main research areas

Internet-of-Things Applications

- How to integrate and manage thousands of devices and sensors?
- How to detect and predict anomalies in large volumes of continuous sensor data?

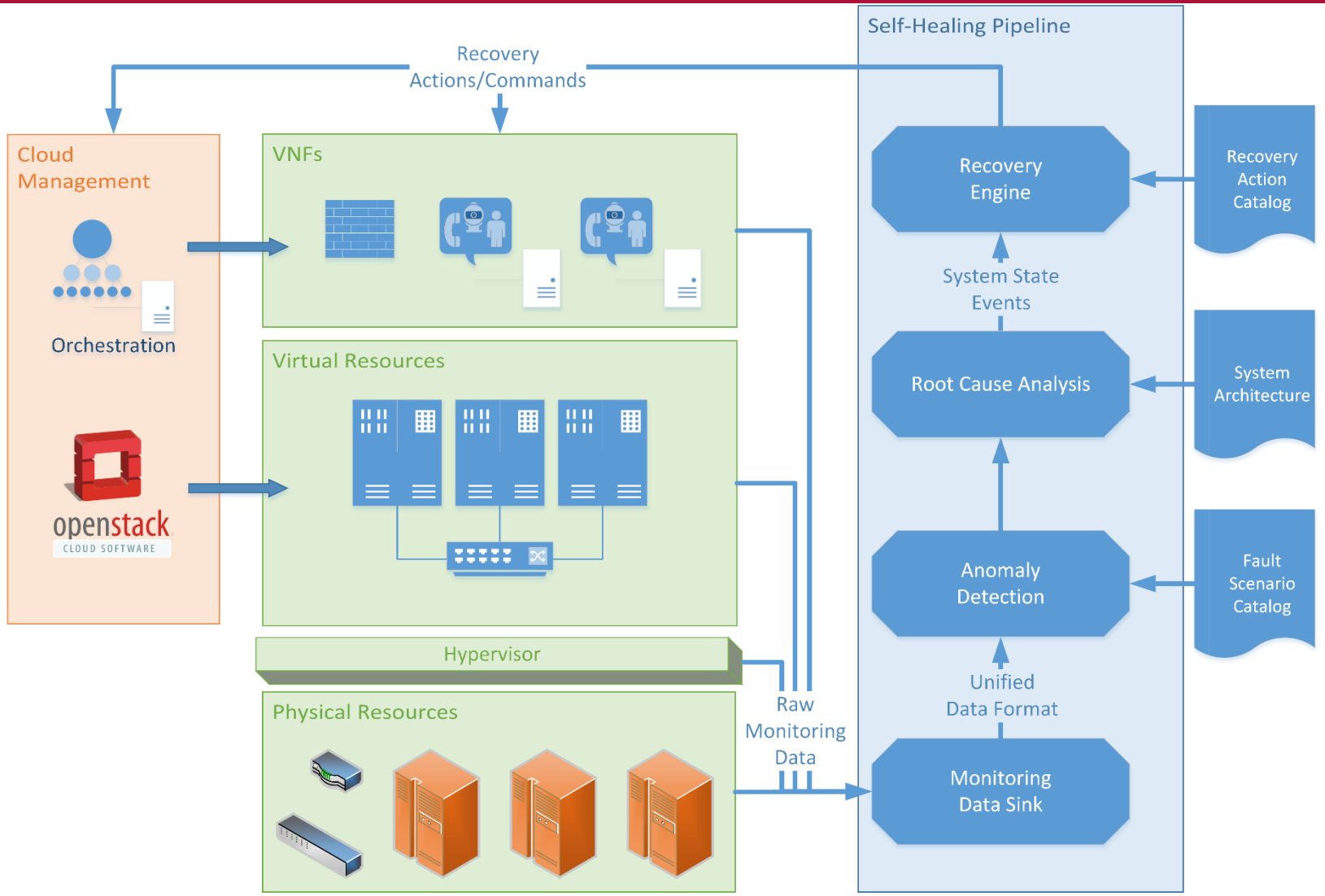
Scalable Adaptive Data-Intensive Middleware

- How to write a program for a thousand computers?
- How to adapt the execution of a job to changing resource requirements?

QoS-Aware Operation of Virtualized Infrastructures

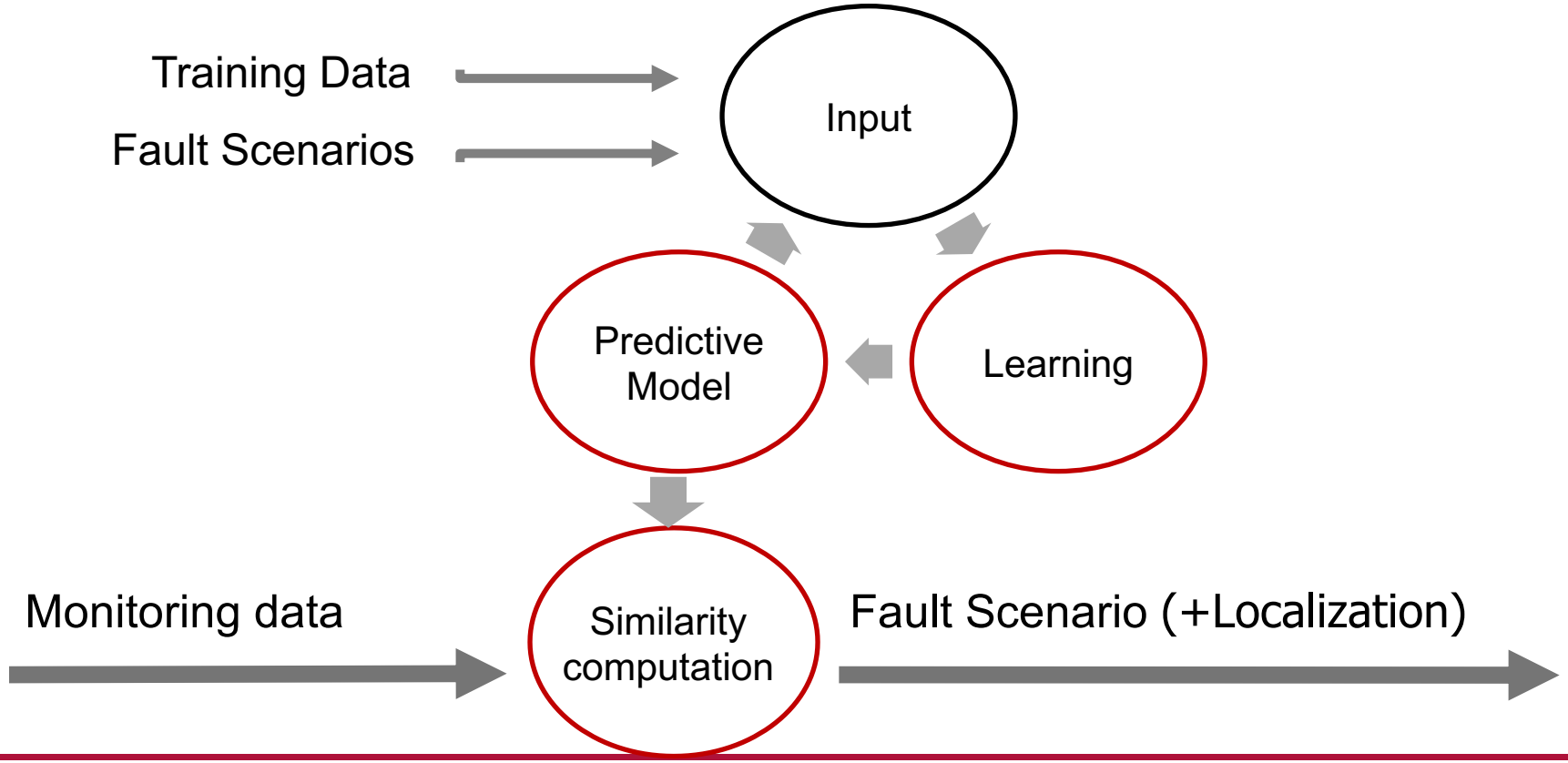
- How to operate a complex and large IT infrastructure?
- How to guarantee the required quality of a service to the customer?

Cloud Recovery Engine

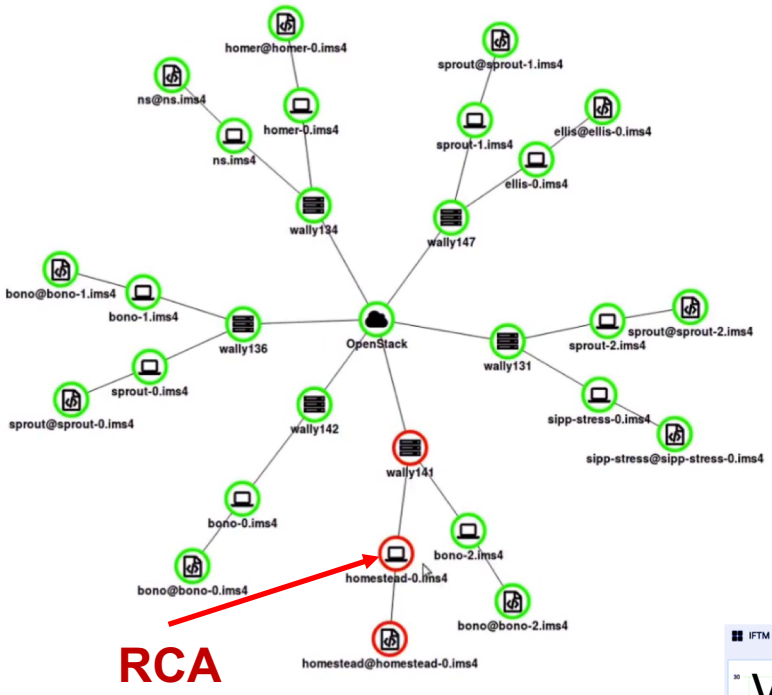


Anomaly Detection

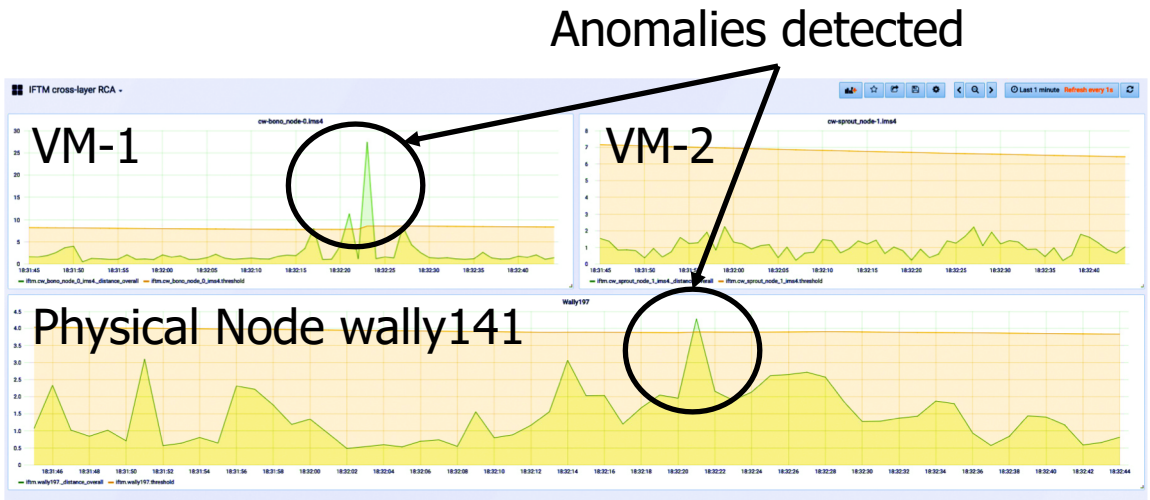
- Anomaly can appear anywhere and with any intensity
- Real-time detection of degraded state faults



Root Cause Analysis

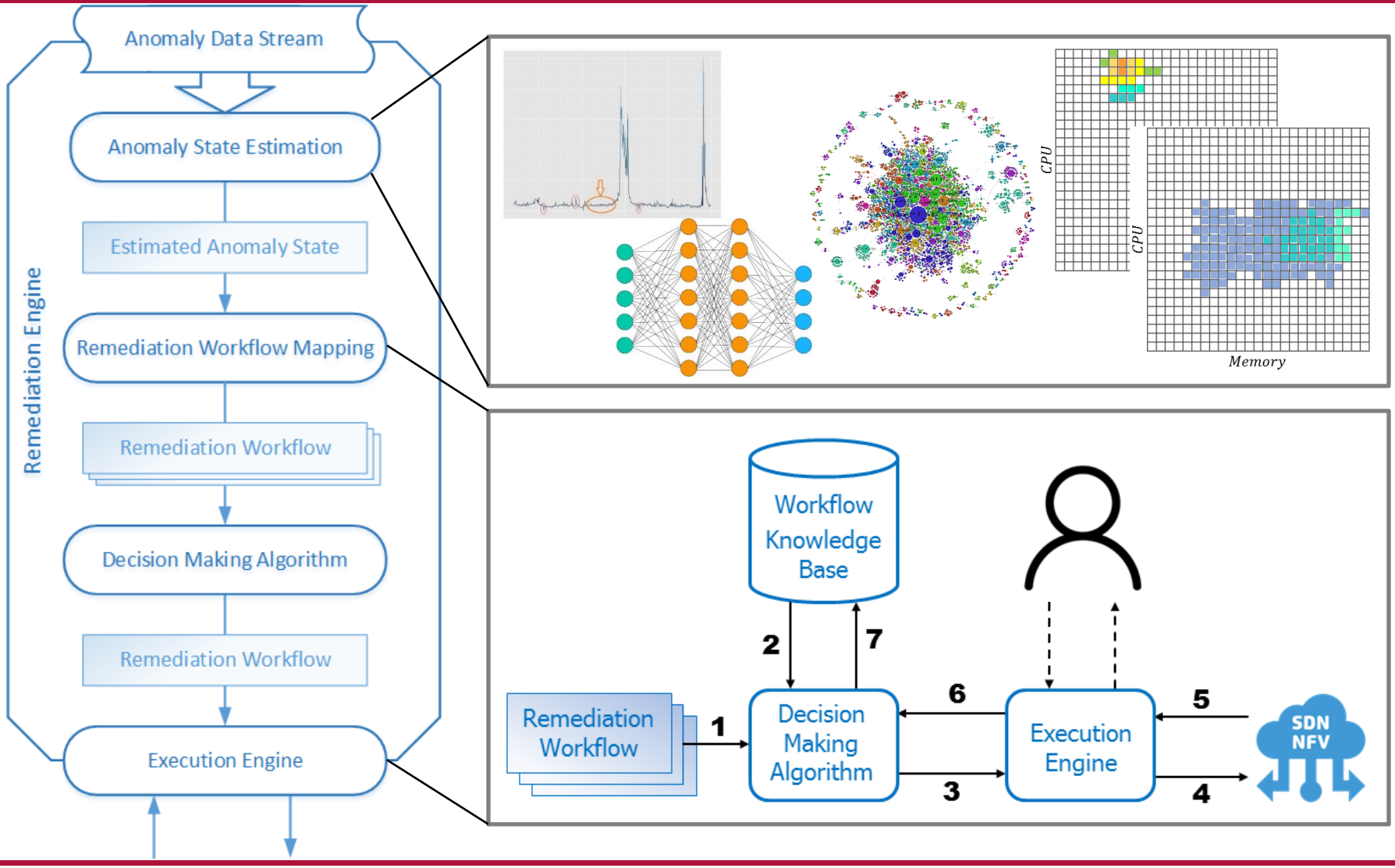


- Several nodes exposing anomalies
- Typical for interdependent component groups
- Aim: Finding the cause of the problem



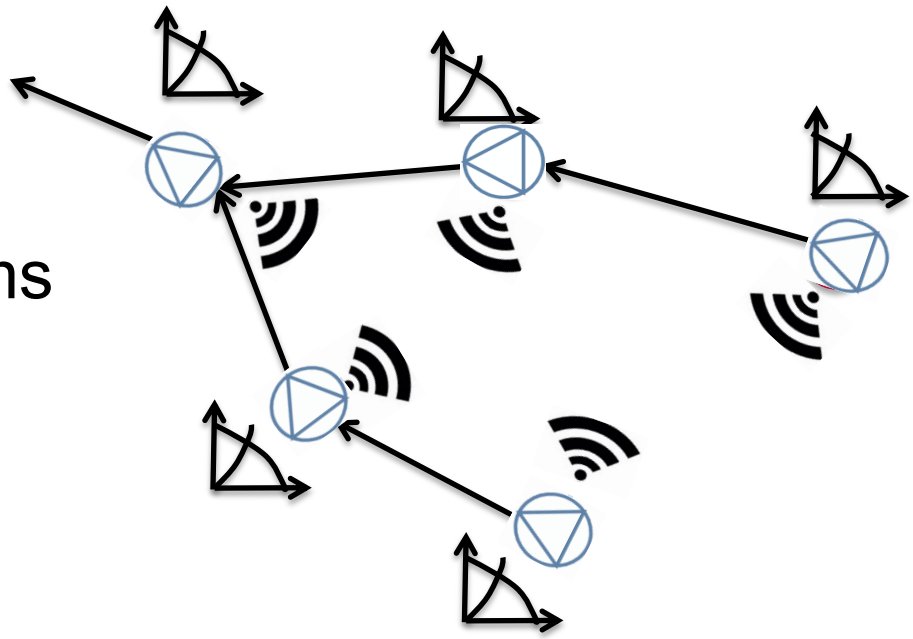
- Current RCA is based on heuristics

Self-Remediation and Recovery



Data-Intensive IoT Applications

- Continuous, reliable, and scalable processing of
- large volumes of sensor data
- for e.g.
  - anomaly detection,
  - load prediction,
  - and optimization
- in cyber-physical systems

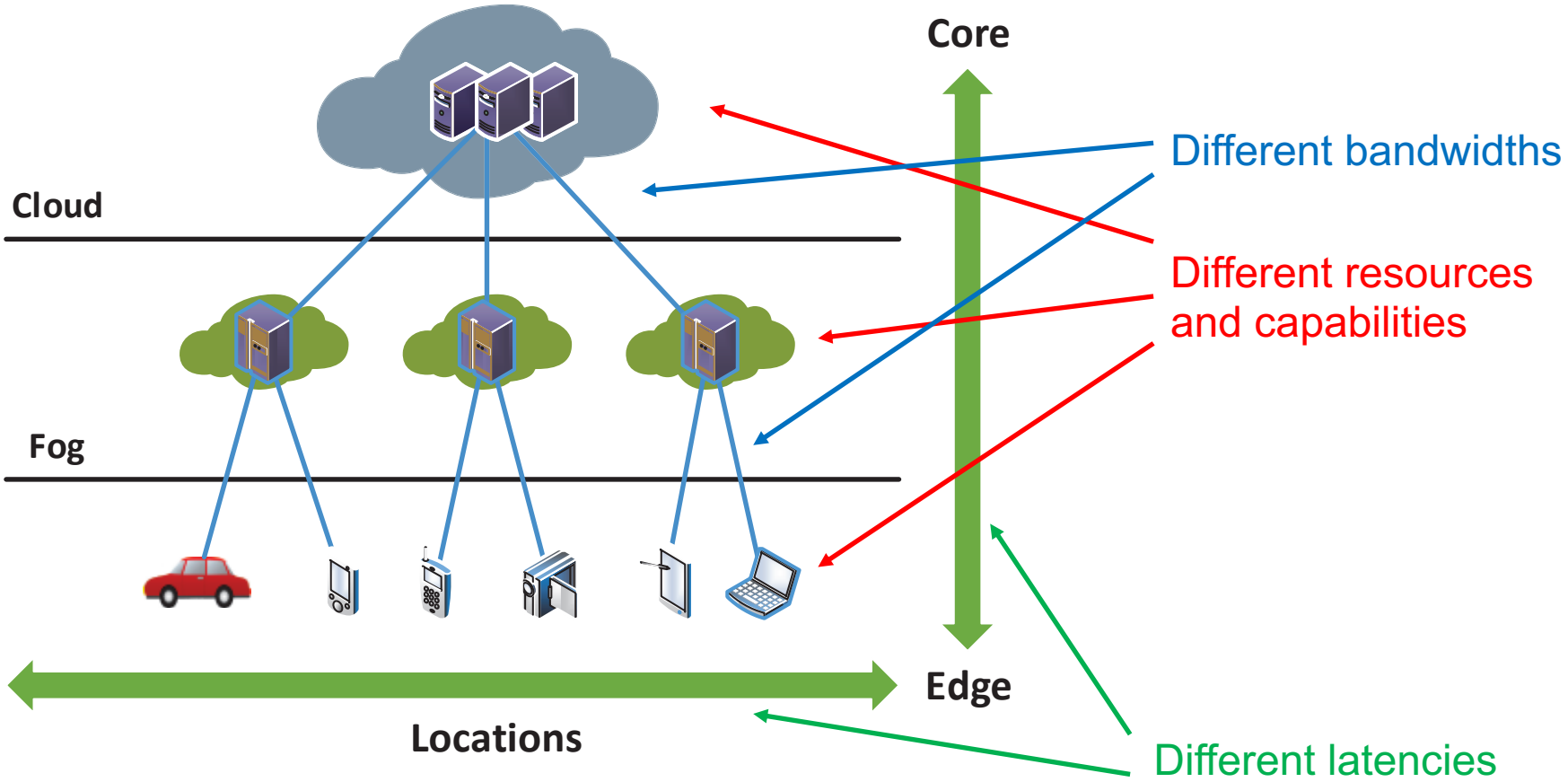


Task Placement in Edge Computing Topologies (1/2)

- IoT increasingly everywhere: Connected Cars, Smart Home, Smart Cities, Industry 4.0

- Typical IoT topologies include
  - Resource heterogeneity: Various connected devices with vastly different resources
  - Geo-distributed device locations: Data is typically generated at many places and then send to a central remote cluster/cloud for analysis (large latencies)

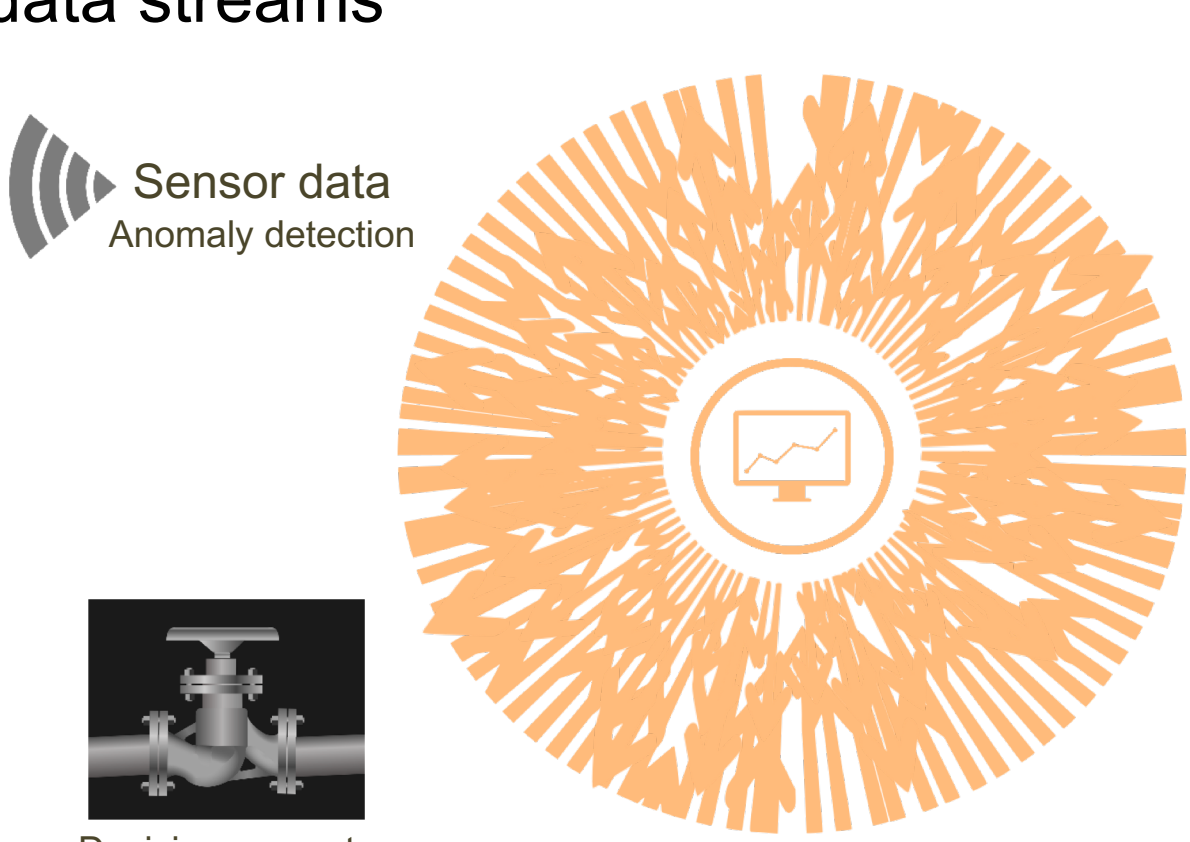
Task Placement in Edge Computing Topologies (2/2)



Ivan Stojmenovic and Sheng Wen. "The Fog Computing Paradigm: Scenarios and Security Issues." *2014 Federated Conference on Computer Science and Information Systems (FedCSIS)*. IEEE. 2014.

Learning from Heterogeneous IoT-Data Sources

- Learning and analytics from heterogeneous data and data streams



References

- [1] P. Mell, T. Grance: “The NIST Definition of Cloud Computing”, Technical Report, National Institute of Standards and Technology, 2011, http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf