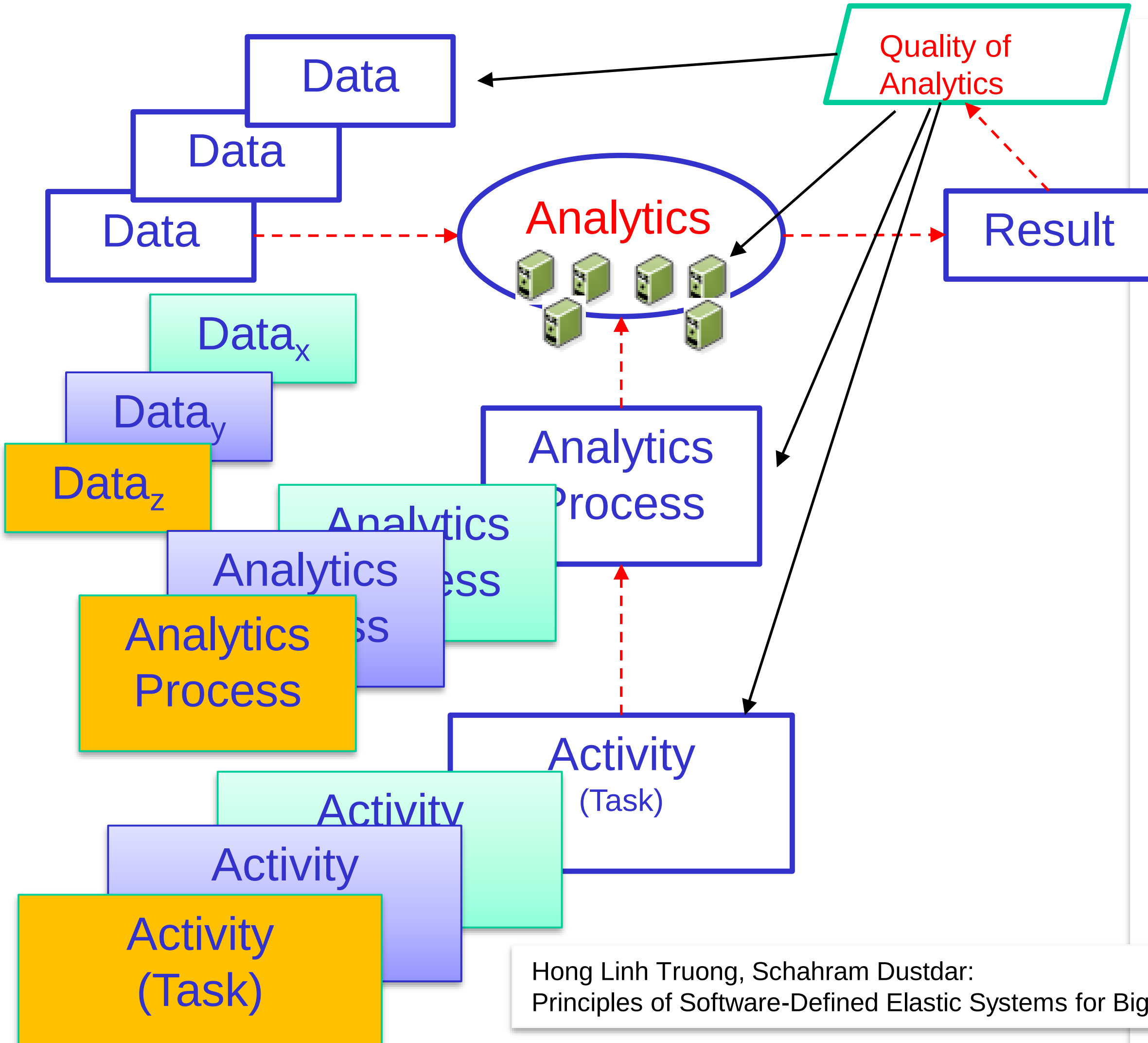
Principles of Elasticity for Service systems

Hong-Linh Truong
Faculty of Informatics, TU Wien

hong-linh.truong@tuwien.ac.at
http://www.infosys.tuwien.ac.at/staff/truong
@linhsolar

Elasticity in (big) data analytics

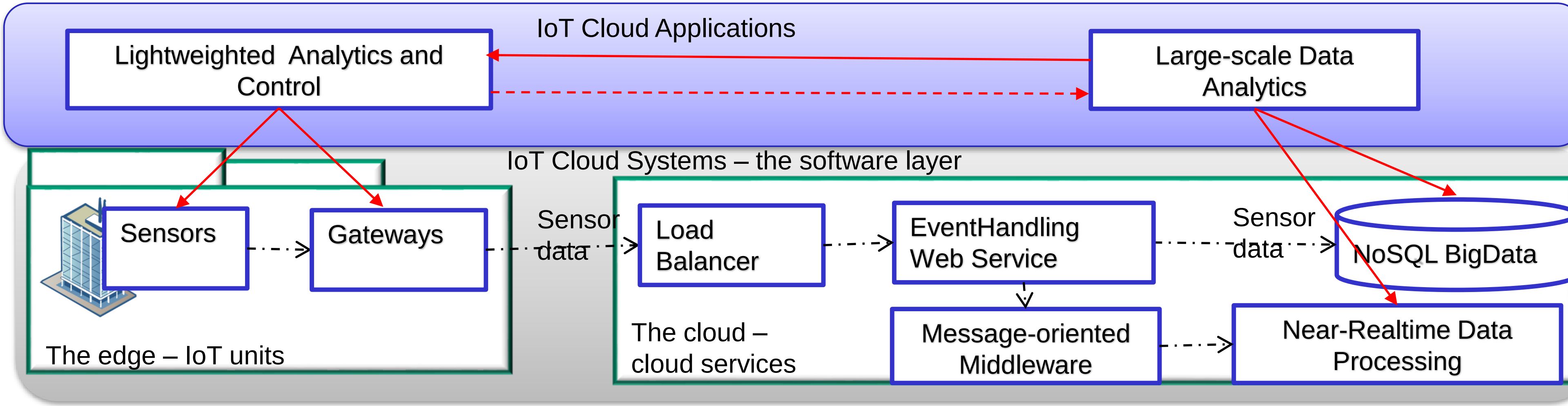


- **More data** → more compute resources (e.g. more VMs)
- **More types of data** → more activities → more analytics processes
- Change **quality of analytics**
  - Change quality of data
  - Change response time
  - Change cost
  - Change types of result (form of the data output, e.g. tree, table, story)

Hong Linh Truong, Schahram Dustdar:
Principles of Software-Defined Elastic Systems for Big Data Analytics. IC2E 2014: 562-567

Elasticity in slices of IoT, Network functions and cloud resources

Application example

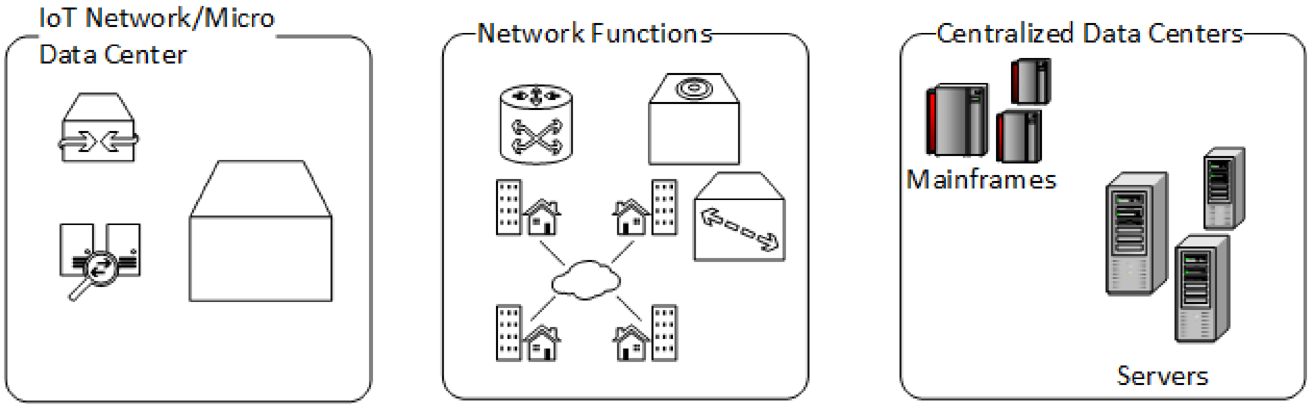


What should we do if suddenly many sensors send a lot of data?

What if you know that “5 minutes from now, 10*n sensors will be started?

Elasticity in slices of IoT, Network functions and cloud resources

„IoT + Network functions + Clouds"



What if in the "network functions" we can create VMs or perform network traffic engineering?

Elasticity principles can be used to support dynamic quality of analytics

Elasticity Principles: Elasticity of data and analysis processes

- Multiple types of objects from different sources with complex dependencies, relevancies, and quality
- Different data and algorithms models for analyzing the same subject
- New analytics subjects can be defined and analytics goals can be changed
- Decide/select/define/compose not only data but also analysis pipelines based on existing ones

Management and modeling of elasticity of data and processes during the analytics

Elasticity Principles: Elasticity of data resources

- Data provided, managed and shared by different providers
- Data associated with different concerns (cost, quality of data, privacy, contract, etc.
- Static data, open data, data-as-a-service, opportunistic data (from sensors and human sensing)
- Distributed big data and multiple data owners

Data resources can be taken into account in an elastic manner: similar to VMs, based on their quality, relevancy, pricing, etc.

Elasticity Principles: Elasticity of humans and software as computing units

- Human in the loop to solve analytics tasks that software cannot do
- Human-based compute units can be scaled up/down with different cost, availability, and performance models
- Human-based compute units + software-based compute units for executing analysis pipelines
- Elasticity controls can be also done by humans

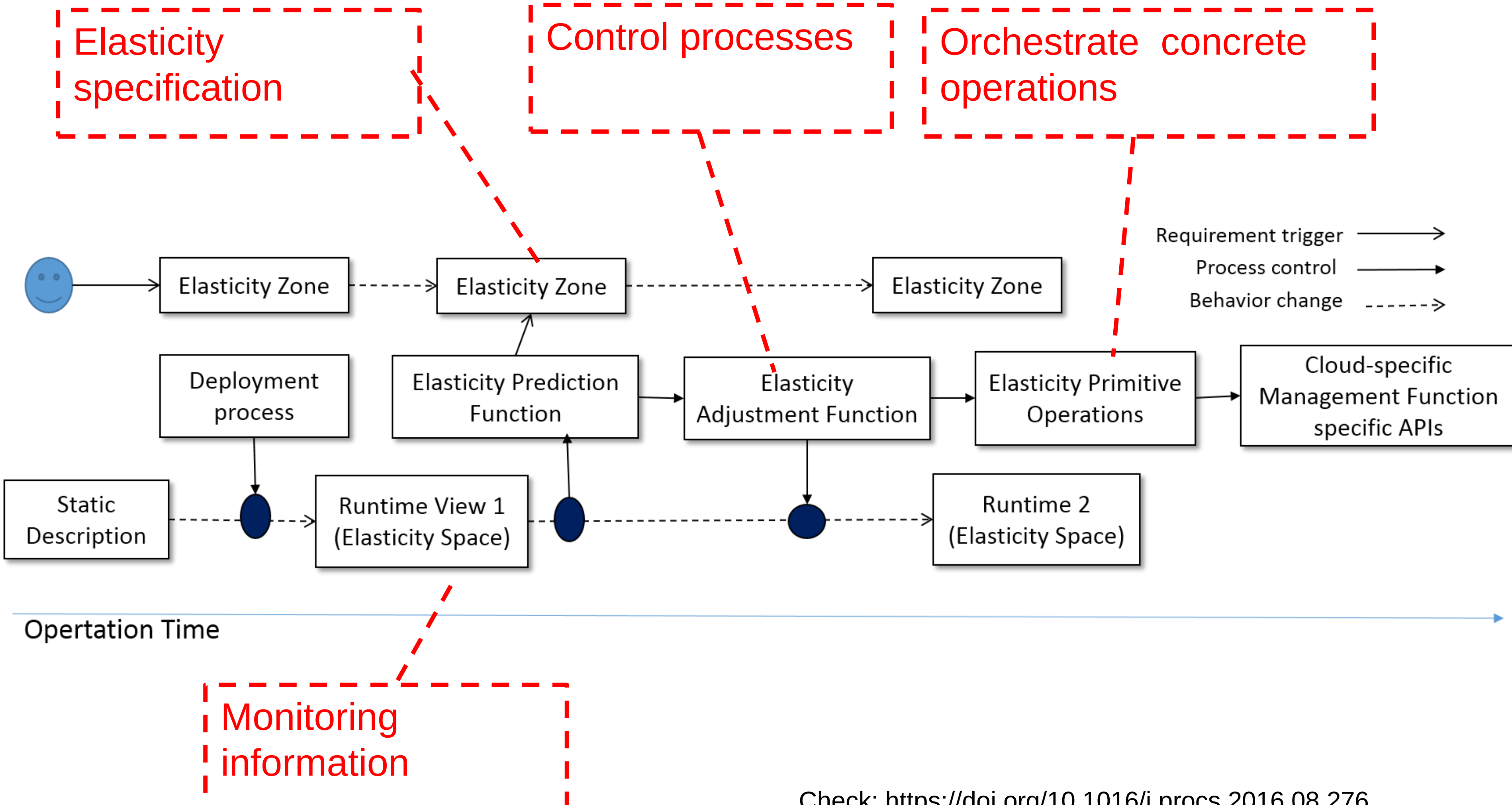
Provisioning hybrid compute units in an elastic way for computing/data/network tasks as well as for monitoring/control tasks in the analytics process

Elasticity Principles: Elasticity of quality of analytics

- Definition of quality of analytics
  - Trade-offs of time, cost, quality of data, forms of output
- Using quality of analytics to select suitable analysis processs, data resources, computing units
- Multi-level control for the elasticity based on quality of analytics

Able to cope with changes in quality of data, performance, cost and types of results at runtime

General software design concept: Lifecycle of applications and elasticity



Check: https://doi.org/10.1016/j.procs.2016.08.276

Exercises

- Read mentioned papers
- Examine possible incidents in your data pipelines
- Examine how QoD evaluators can be integrated into different programming models for QoA-aware data analytics workflows
- Implement some QoD evaluators
- Develop techniques for determining places where QoD evaluators can be performed in your mini projects
- Support data elasticity management in your mini project

Thanks for your attention

Hong-Linh Truong
Faculty of Informatics, TU Wien
hong-linh.truong@tuwien.ac.at
http://www.infosys.tuwien.ac.at/staff/truong
@linhsolar