VALID'2016

www.huawei.com

HUAWEI TECHNOLOGIES CO., LTD.

Panel: Systems & Data – Challenges in Simulation and Verification/Validation for Complex Systems & Big-Data

Moderator

Xinli Gu, Huawei Technologies, USA

Panelists

- Jos van Rooyen, Indentify Test Services, the Nederland
- Martin Zinner, Technische Universitat Dresden, Germany
- Yingzhen Qu, Huawei Technologies, USA
- Simon Genser, Kompetenzzentrum das Virtuelle Fahrzeug, Austria
- Volker Gollucke, OFFIS Institute for Information Technology, Germany

Thank You

www.huawei.com

Derivative Estimation via DELC-Method with Noisy Data

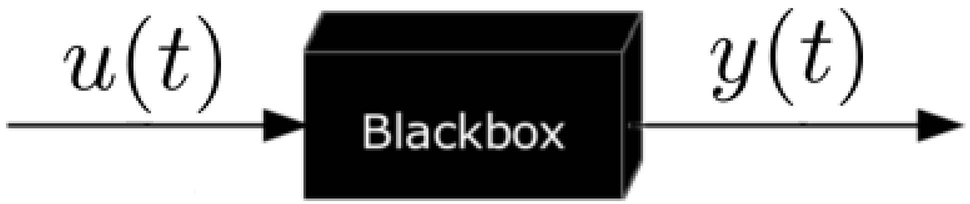
Simon Genser
VIRTUAL VEHICLE Research Center

VIRTUAL VEHICLE Research Center is funded within the COMET – Competence Centers for Excellent Technologies – programme by the Austrian Federal Ministry for Transport, Innovation and Technology (BMVIT), the Federal Ministry of Science, Research and Economy (BMWFW), the Austrian Research Promotion Agency (FFG), the province of Styria and the Styrian Business Promotion Agency (SFG). The COMET programme is administrated by FFG.

My Person

- **Simon Genser**
- **Completed Master in Industrial Mathematics at the TU Graz, Austria.**
  - numerical methods (FEM, BEM, FDM,…)
  - functional analysis
  - partial differential equations
- **Employed at the Virtual Vehicle Competence Center in Graz, Austria.**
  - Junior Researcher
  - working on my PhD thesis about coupling of control systems
- **Fields of Interest:**
  - numerical derivative estimation
  - coupling of control systems

Example: Sensitivity Analysis of a Blackbox System



Interested in Output-Input Sensitivity: $\dfrac{dy(t)}{du(t)}$

Computation:

$$\frac{dy(u(t))}{dt} = \frac{dy}{du}\frac{du}{dt} \quad \Longleftrightarrow \quad \frac{dy}{du} = \frac{dy}{dt} \Big/ \frac{du}{dt}$$

So we need:

$$\boxed{\frac{dy}{dt}, \frac{du}{dt}}$$

Computation formula:

$$y_k^{(j)} = \frac{1}{h^j} \sum_{i=0}^{N} c_{i_j} y_{k-i}$$

First solve:

$$A\boldsymbol{c} = \boldsymbol{b}$$

$$A[i, l] := (-\, l)^i$$

$$\boldsymbol{c}[l] := c_{l_j}$$

$$\boldsymbol{b}[i] := \begin{cases} i! & \text{for } i = j \\ 0 & \text{otherwise} \end{cases}$$

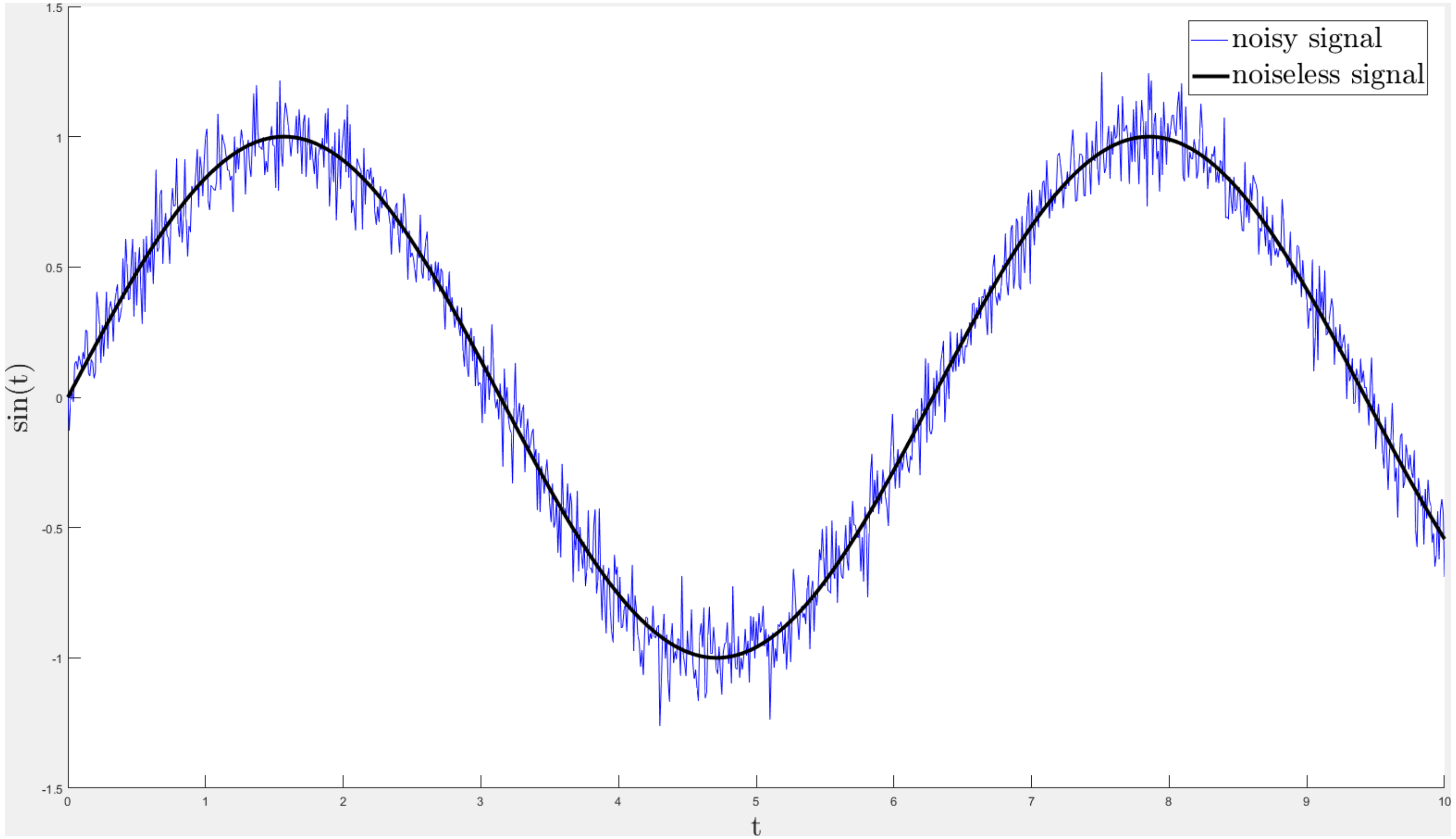
One is interested in $\dot{y}_k$ only by samples $y_k, \ldots, y_{k-N}$.

Therefore use the DELC method

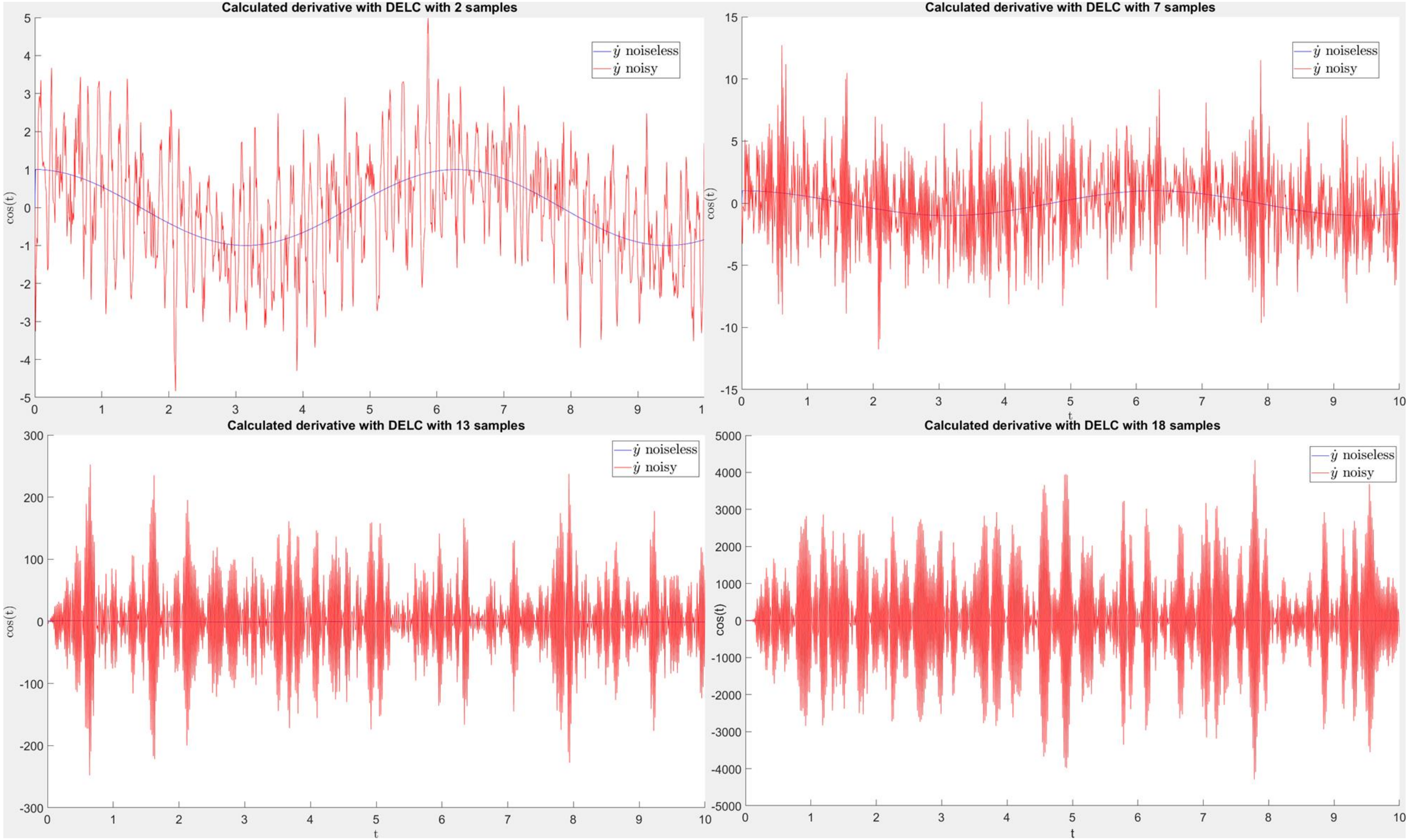
$$\dot{y}_k = \frac{1}{h} \sum_{i=0}^{N} c_i y_{k-i}$$

For noiseless signals it works

but for noisy signals it fails.



DELC with noisy signal



1. **You can see the DELC Method as interpolating the samples $y_k, \dots, y_{k-N}$ with a polynomial and derive this polynomial to get the derivative $\dot{y}_k$.**

2. **Affine Arithmetic Theory:**

   Noisy samples are represented as

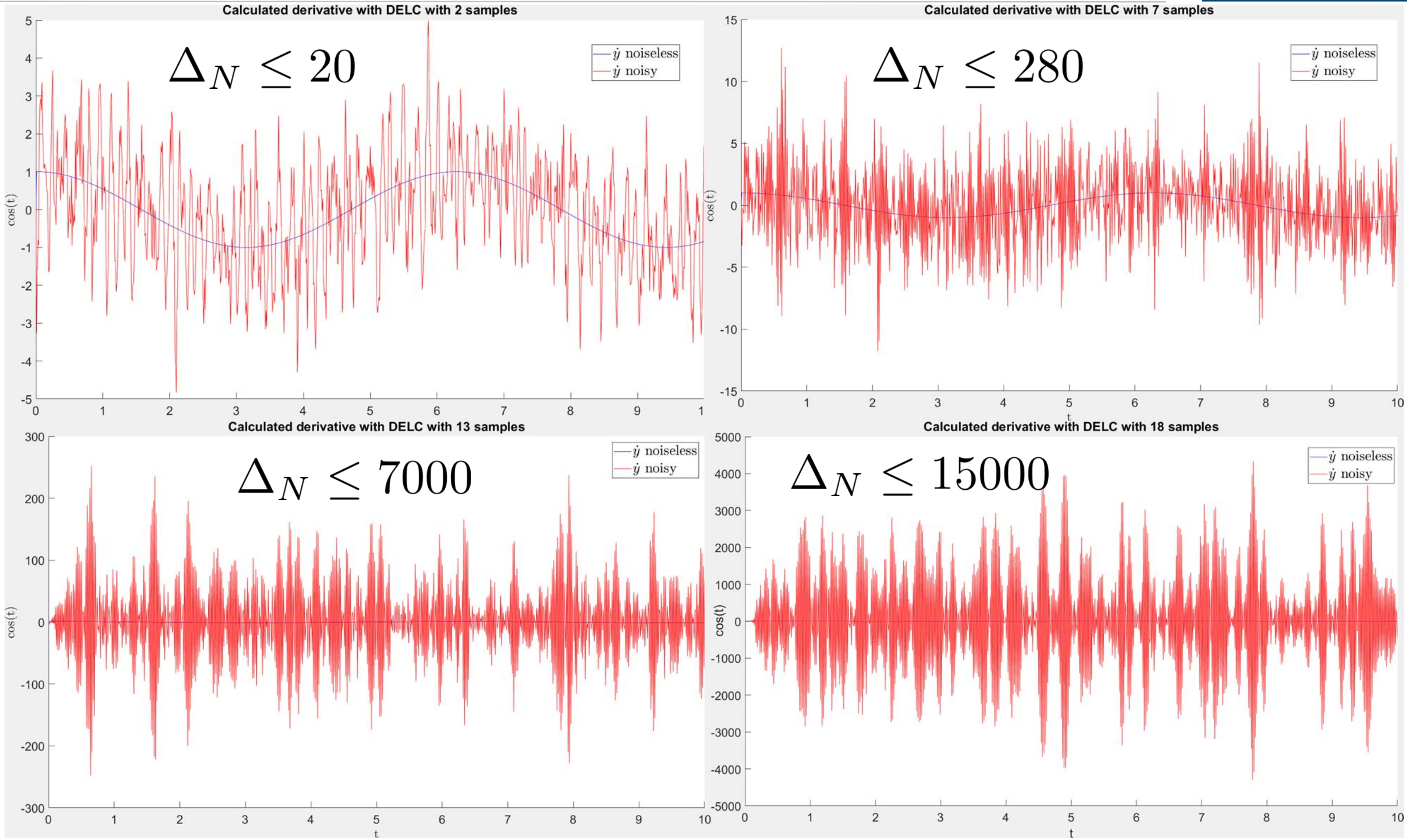
$$y_{k-i} = y_{0_{k-i}} + \epsilon_i \cdot \widetilde{y_{k-i}} \quad \epsilon_i \in [-1, 1]$$

   with $\widetilde{y_{k-i}}$ as the amplitude of the white noise.

   The uncertainty Δ_N in the DELC computation is

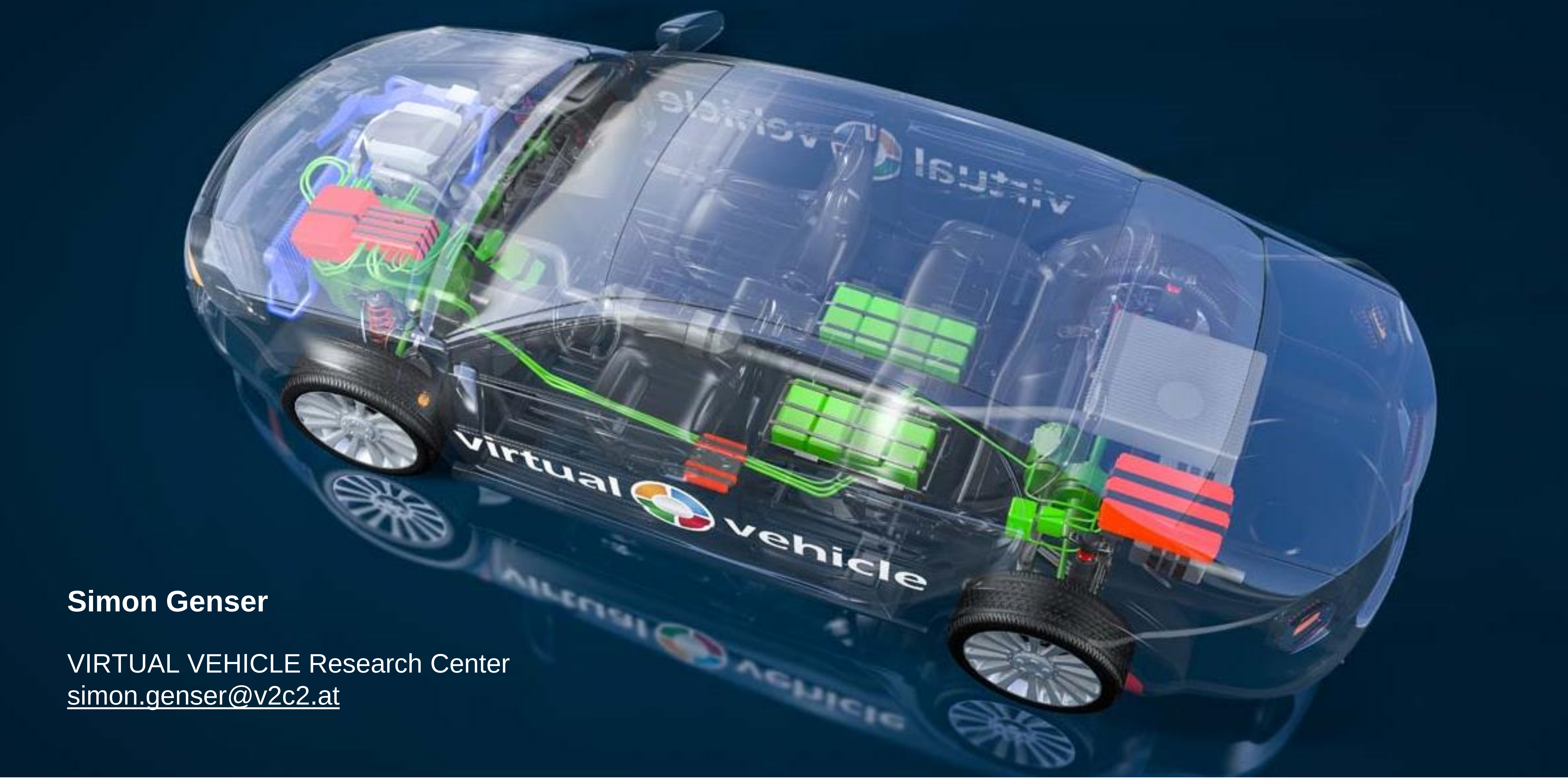
$$\Delta_N \leq \frac{1}{h} \sum_{i=0}^{N} |c_i \cdot \widetilde{y_{k-i}}|$$

Affine Arithmetic and DELC with noisy signal



1. **Smooth the signal first and then use DELC on the smoothed signal**
   1. Smooth with Least Square Methods
   2. ….

2. **Some adjustment to the DELC to get rid of the interpolating behaviour**
   1. Try something like approximation instead of interpolation
   2. Other ideas=?



Simon Genser

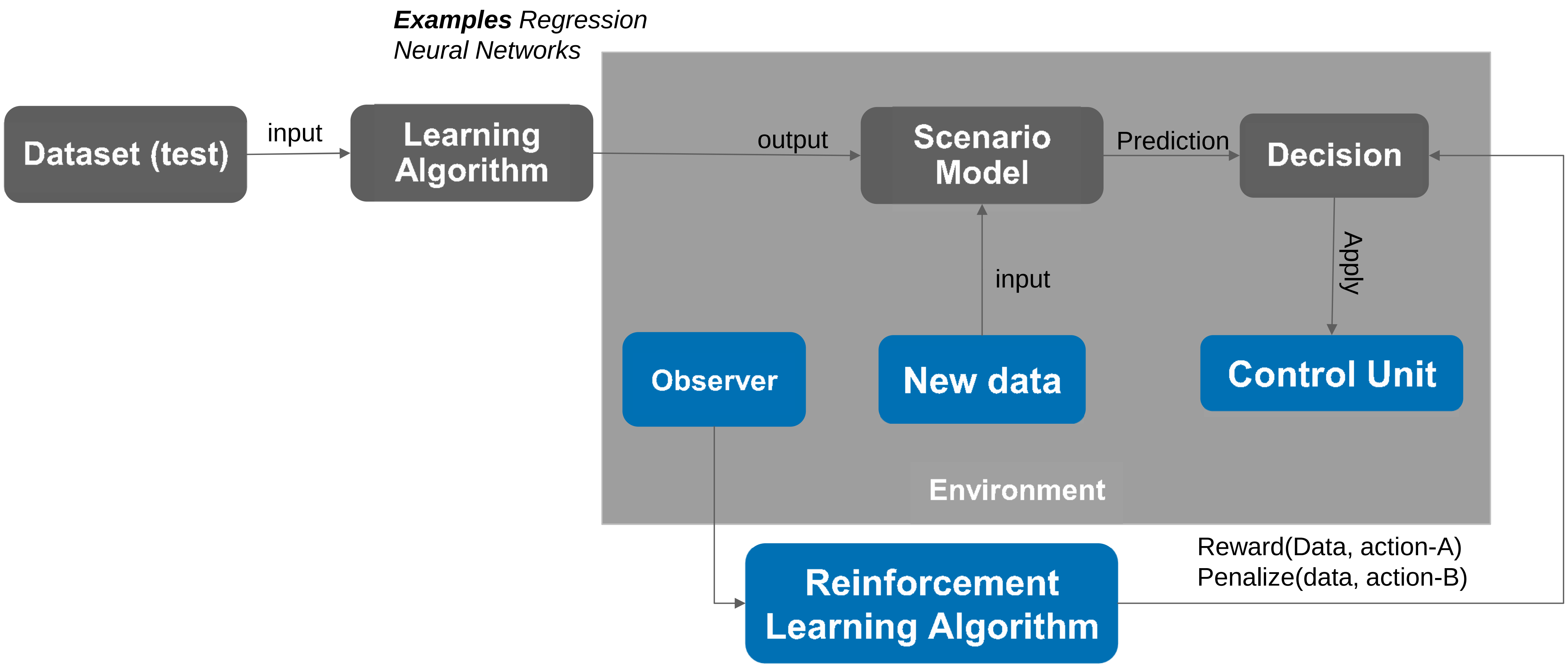
VIRTUAL VEHICLE Research Center
simon.genser@v2c2.at

Intelligent Data Center with Machine Learning

Yingzhen Qu
Senior Research Scientist, Future Networks
Huawei Technologies, US Research Center

October 2017

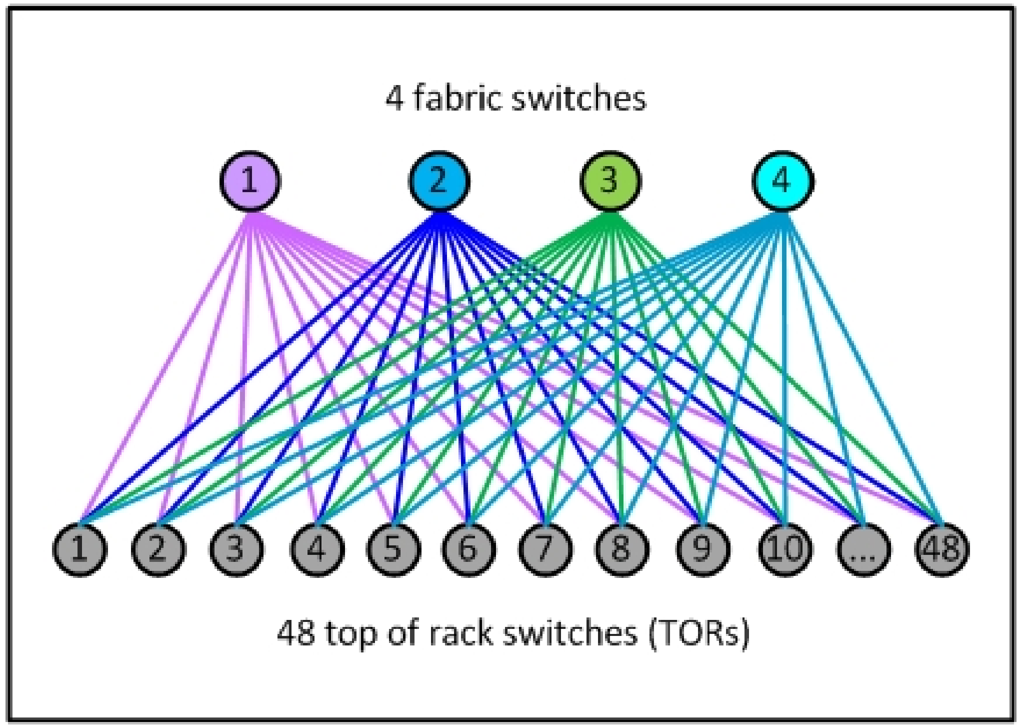
A Generalized Machine Learning Loop

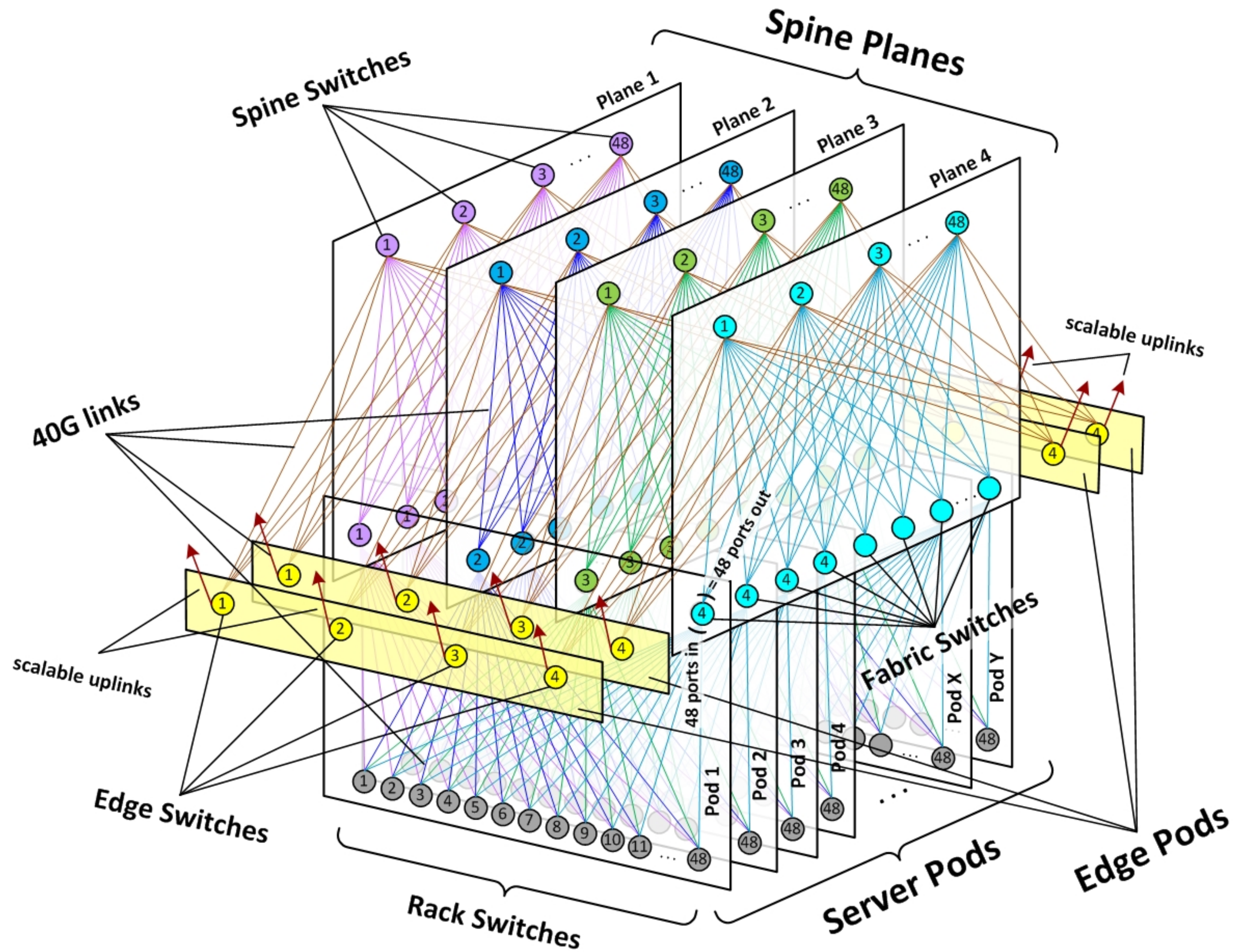


An Example: ECMP Based Link Utilization Problem in a Switch

- **Massive Scale DCs use fixed spine-leaf topology**
- **ECMP distributes traffic across multiple paths**
- **ECMP uses Hash computation to balance similar flows over multiple links**
- **However, the flows are not evenly balanced**
  - Low-bandwidth (Mice) flows: Majority of flows are short-lived and latency sensitive.
    - Example: Web, chat applications
  - High-bandwidth (Elephant) flows consume majority bandwidth and are long-lived.
    - Example Storage-intensive big-data, data-replication and backup applications
- **Problem**
  - Variance in the amount of bandwidth used between long-lived vs short-lived flows does not ensure that traffic is balanced across all the links.
  - Increase in Mean-time-to completion for mice flows
  - Reduced data-rate for elephant flows due to congestion control

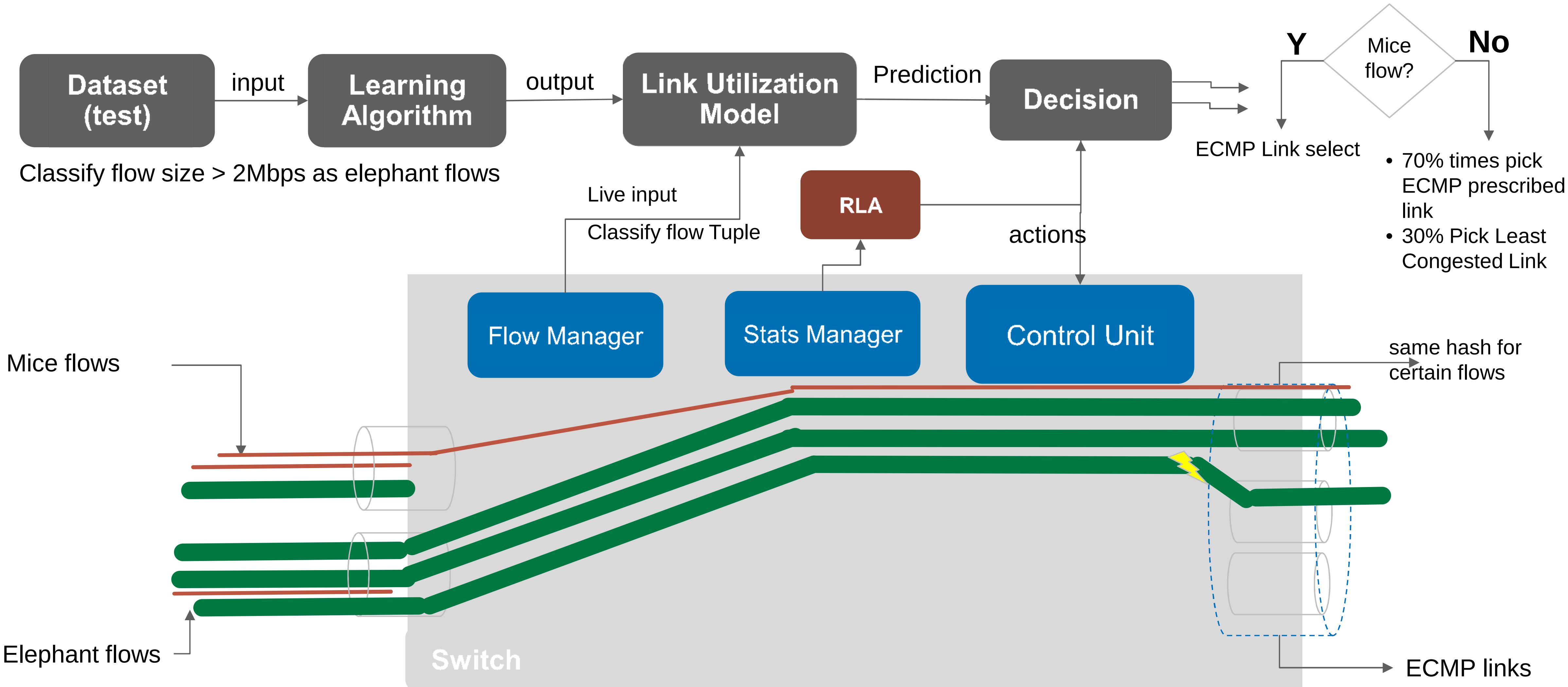
Facebook Data Center Fabric





Source: https://code.facebook.com/posts/360346274145943/introducing-data-center-fabric-the-next-generation-facebook-data-center-network/

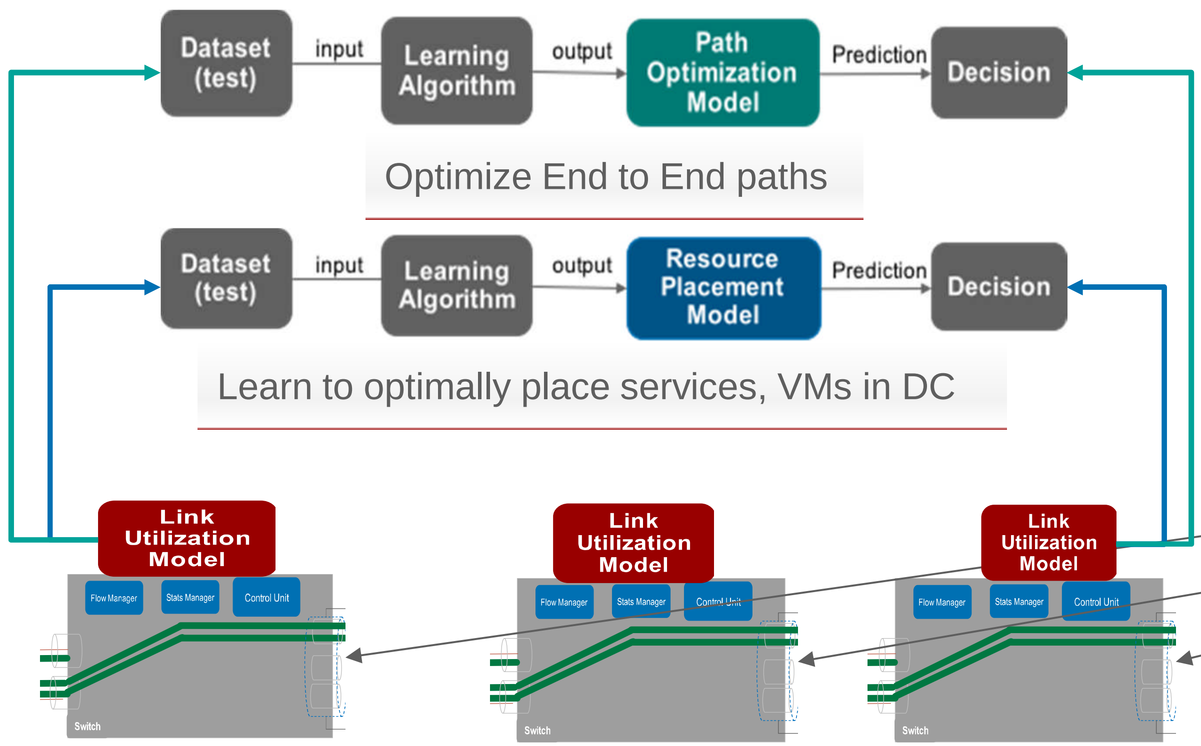
Machine Learning for ECMP Link Utilization in a Switch

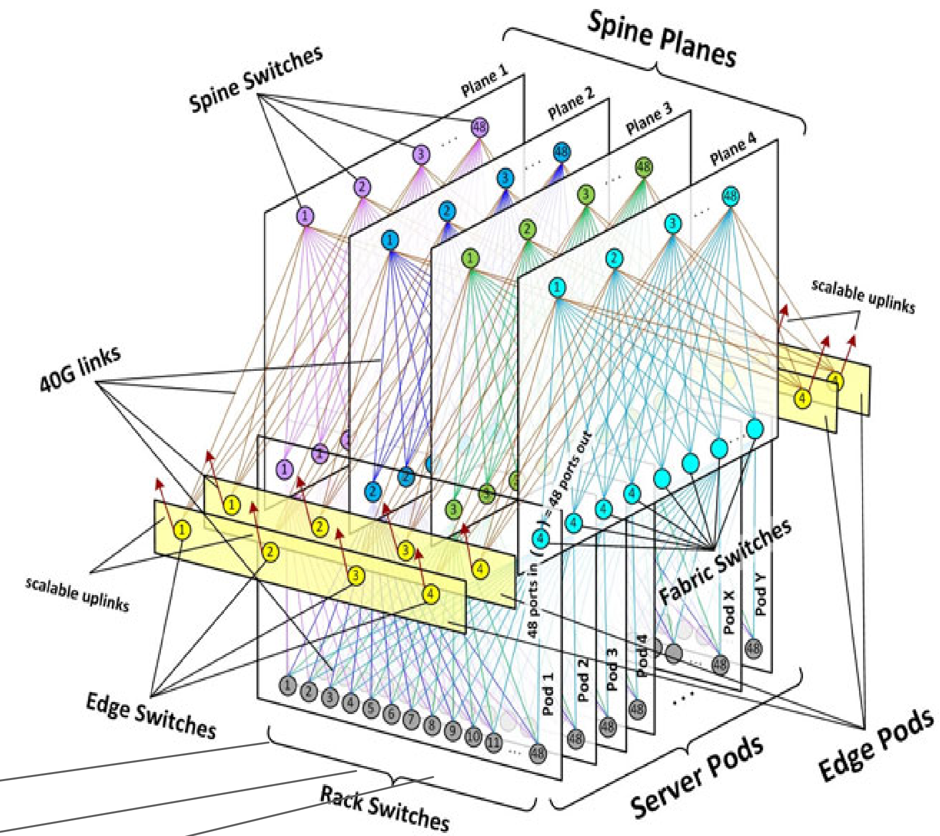


RLA – Reinforcement Learning Algorithm

Intelligence Driven Networking – DC Scenarios with Global Scope

- Extend to wider scoped learning - Global models across multiple switches
- Different Learning models for different scenarios together





Src: Facebook data center

IT mediation

Jos van Rooyen

Who am I?

- Employed at Identify as partner / principal consultant
- >25 years in software testing & quality management
- Co-author TestGrip ,TestFrame, Project de Baas, Quality Supervision, Textbook; “Aan de slag met software testen”
- Test expert online magazine Computable
- Publication areas; Test process Improvement, BI-testing, Test automation, Test Education, Risk Based Testing, Quality Supervision
- Review committee software conference Valid2017
- Founder of the “Houten Groep”
- Member NESMA working party; Metrics in Contracts
- Member of several working parties Dutch Testing Society:
  - Model Based Testing
  - Test Education universities of applied science

Interpretation of the problem

Contracts by suppliers are very promising

A lot of functionality for a low price

I see a lot of projects failing just at this point:

- Wrong functionality or less functionality delivered
- Misinterpretation of the requirements
- Delivery of the system too late
- Budget problems. Running out of money
- Poor quality so the demanding organization has to do a lot of rework
- Demand / supply doesn't work together

Possible solutions to solve the problem

Based on IT mediation:

Determination of the suitability of the new information system in the existing organization, with respect to the business process, the people, IT and the organization:

- Quality
- Maintainability
- Metrics
- Process
- Vendor management
- Contract management
- Legal

4

Possible solutions to solve the problem(2)

- Based on metrics:
  - Number of defects discovered
  - Defect Removal Efficiency
  - Compliance to the requested functional quality
  - Quality in use
  - Coverage degree requirements

5

Questions to the audience

- Do you recognize the described situation?
- What are your ideas to manage these kind of situations?
- Are the described solutions indeed good solutions?

6

Questions?

Jos van Rooyen
E-mail: jos.van.rooyen@identify.nl
Mobile: +31-654906282
LinkedIn: Jos van Rooyen

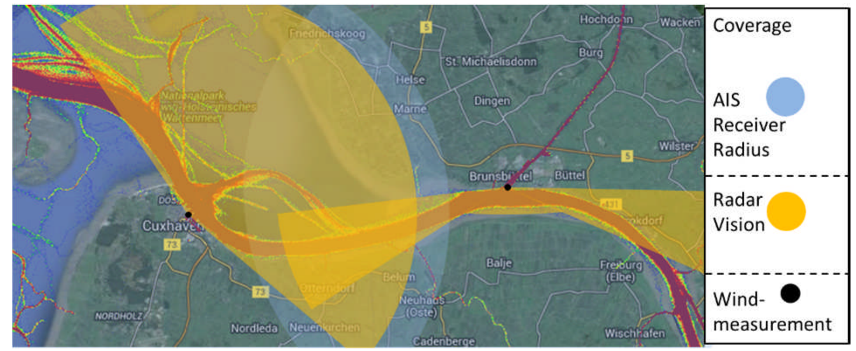
Data Analysis for a Reference Waterway


Observe, , Test

Dr.-Ing. Volker Gollücke

Reference Waterway

The Reference Waterway covers the Elbe and Kiel Canal Approach, Germany. It covers a basic maritime surveillance infrastructure with three Sensorboxes (including AIS, Radar, cameras) and broad band communication.

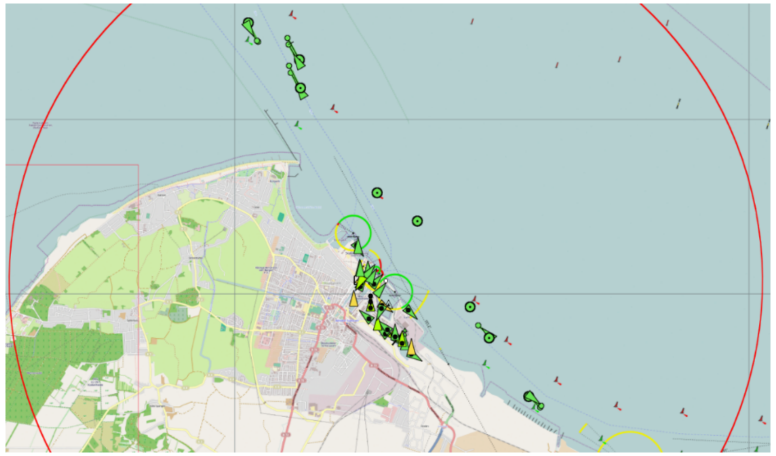


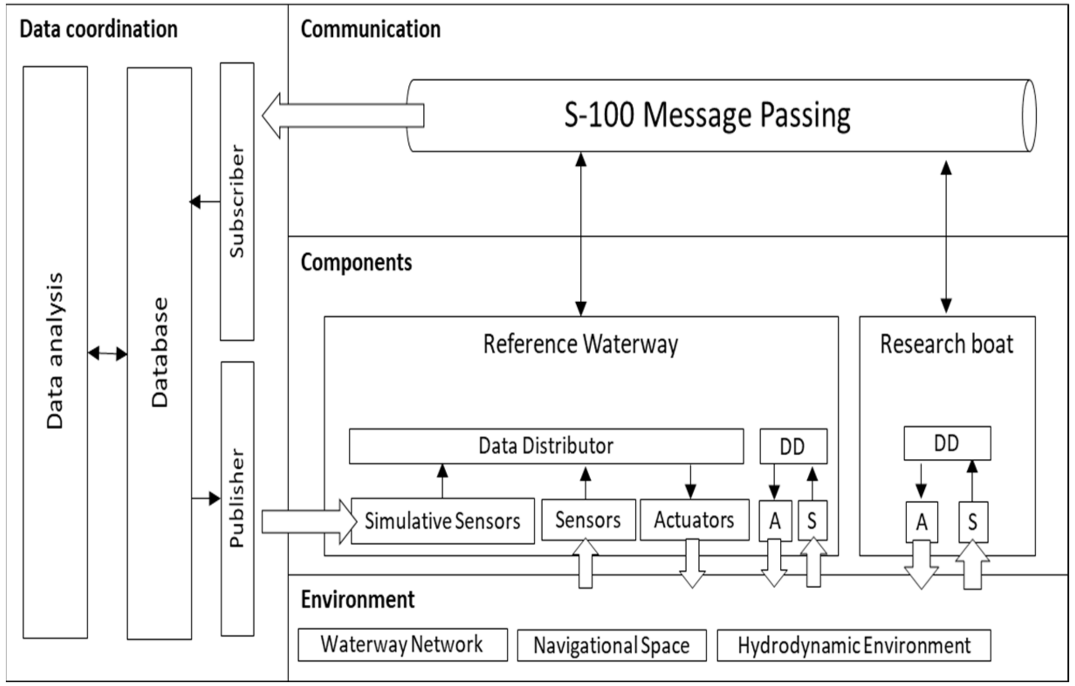
- Setting up a database with travel pattern and near collisions
- Gathering Live Radar, AIS, Wind and Video Data
- Connected to the eMir Infrastructure with LTE

Observing and learning from the real world

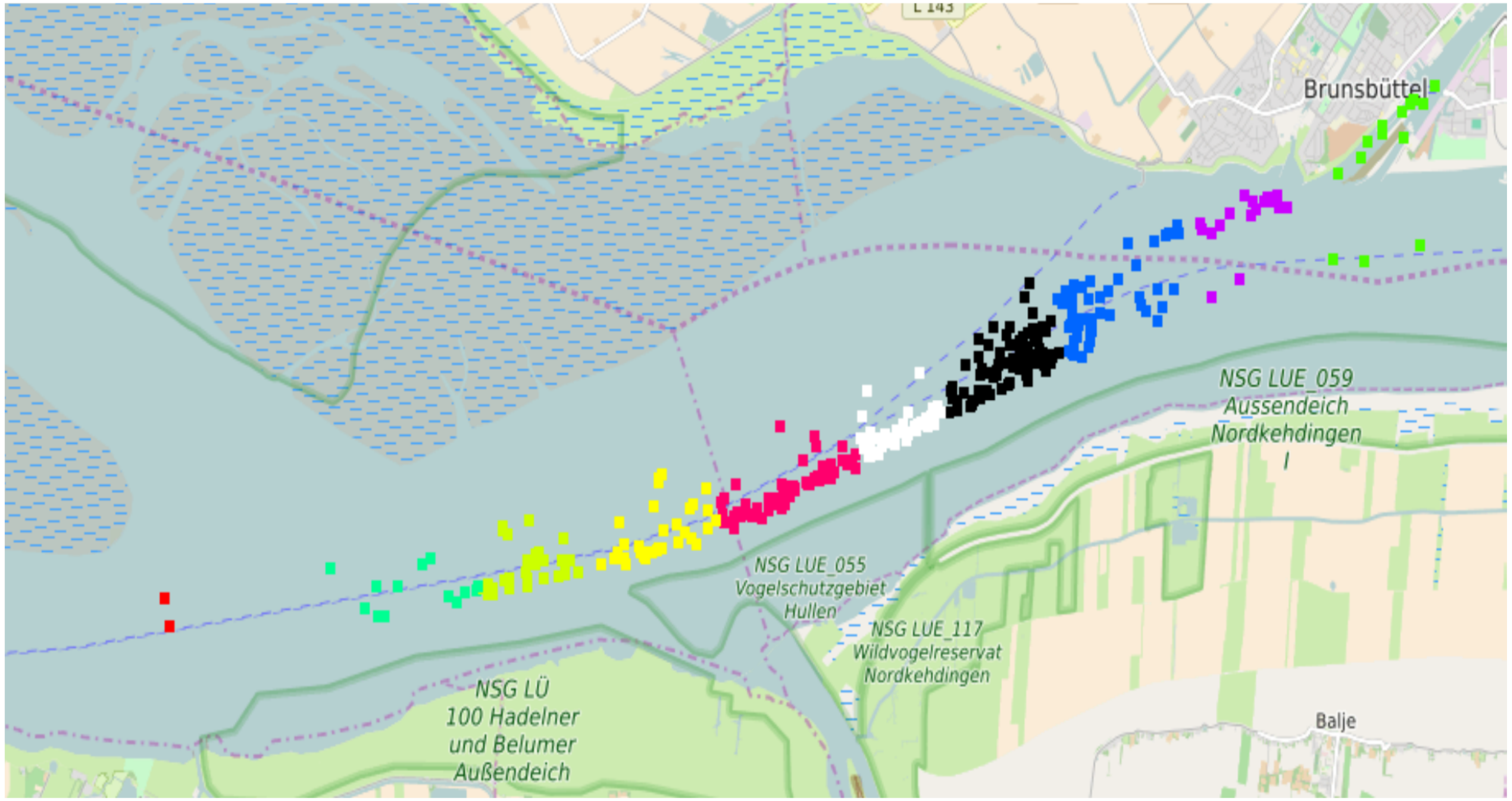
All sensors can be gathered for an overview of the situation on the Elbe

eMir Infrastructure allows data exchange between different sources and users



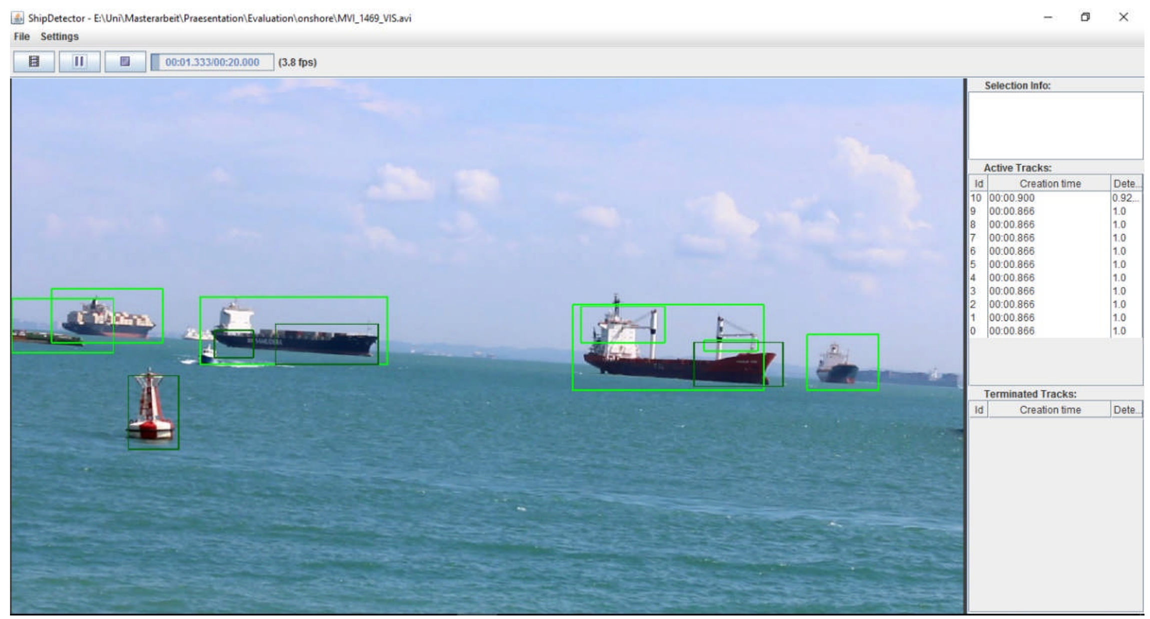


Example: Maneuverpoints



application of the knowledge acquired

- Research Vessel
- Not only AIS data but also Video Streams
- Online prediction of behavior of target vessels



Manufacturing Analytics

Methods to analyze and improve the performance of a fab

Martin Zinner

October, 10 2017 @ ICSEA 2017 (Athens)

Agenda

Challenges

Little's Law

Operating Curve Management

Conclusion

Literature

Agenda

Challenges

Little's Law

Operating Curve Management

Conclusion

Literature

Challenges in a fab

- Expensive tools should not lack of material to process.
- Tools must be maintained having downtimes (planned and unplanned).

Technical Objectives:

1) Reduce inventory and implicitly reduce cycle time.
2) Increase throughput.

Agenda

Challenges

Little's Law

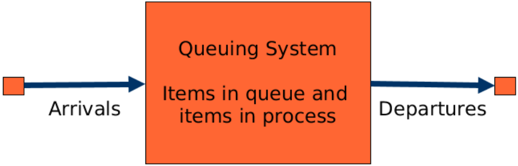
Operating Curve Management

Conclusion

Literature

Little's Law

- **Average cycle time** ("flow time" / "sojourn time")
  $(avg)CT$: Time interval that a unit spent in the system
- **Average work in process - inventory**
  $(avg)WIP$: Average number of items in the system
- **Throughput**
  TH: Average number of items arriving per unit time



The average departure rate exists and equals the average arrival rate:

$$(avg)CT = \frac{(avg)WIP}{TH}.$$

Little's Law: Valability Conditions

Valability Conditions:

- In steady state and non preemptive systems (properties are unchanged in time, no interrupts and later resumes)
- In stochastic systems if the probabilities that some events occur in the system are constant
- Valid as limit on time (holds eventually)

Properties:

- Average departure rate exists, equals the average arrival rate.
- It can be used on all level of the production system (tools, single production step, group of production steps, whole production line for a product, whole production line).
- It is one of the central theorems of the operational research.

Agenda

Challenges

Little's Law

Operating Curve Management

Conclusion

Literature

Flow Factor

Flow factor FF is a reference value, controls the production.

- **Cycle time** CT: Time to process unit including waiting time
- **Raw process time** RPT: Minimal time to process a unit

$$FF := \frac{CT}{RPT}.$$

FF can be calculated by ...

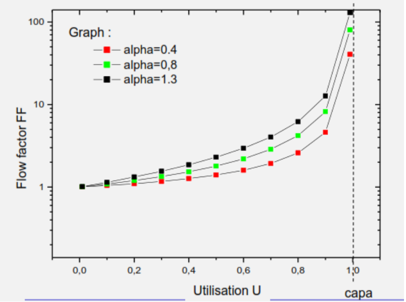
a) ...applying Little's Law to estimate the cycle time,

b) ...using the effective values of the cycle time.

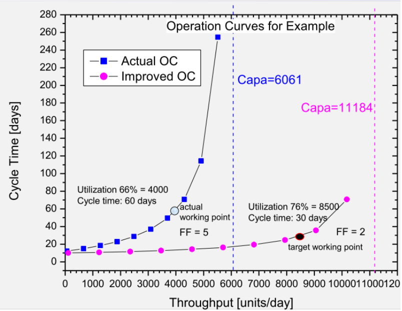
Capacity, Utilization, Variability [3]

- **Capacity** *Capa*: Maximal throughput
- **Utilization** U: Throughput / Capacity
- **Variability** α: Can be calculated using a single (FF, U) tuple

$$FF = \alpha \cdot \frac{U}{U-1} + 1.$$



Utilization vs. flow factor



Throughput vs. cycle time

Agenda

Challenges

Little's Law

Operating Curve Management

Conclusion

Literature

Conclusion

Requested optimization objectives:

1) Reduce inventory and implicitly reduce the cycle time.
2) Increase throughput.

Conclusion:
The two objectives are contradictory!
They cannot be optimized at the same time!

Challenge:
Seeking a good balance between both objectives.

Agenda

Challenges

Little’s Law

Operating Curve Management

Conclusion

Literature

Literature

1) **Book, one of the best:**
   Wallace J. Hopp, Marc L. Spearman: “Factory Physics: Foundations of Manufacturing Management”, 2nd edition, 2001, McGraw-Hill, Higher Education
2) **PhD thesis, easy to follow:**
   Kirsten Hilsenbeck: “Optimierungsmodelle in der Halbleiterproduktionstechnik”, 2005, Doctoral dissertation, Technische Universität München,
   `https://mediatum.ub.tum.de/doc/601620/601620.pdf`
3) **Slides, very accurate and comprehensive:**
   Walter Hansch, Tina Kubot: ‘Factory Dynamics”, Chapter 7, Lectures at Universität der Bunderwehr Munich,
   `http://www.unibw.de/rz/dokumente/fakultaeten/getFILE?fid=6186607&tid=fakultaeten`

Real-Time Data Analytics

Martin Zinner

October, 10 2017 @ ICSEA 2017 (Athens)

Agenda

Introduction

Practical Example

Summary

Bibliography

Agenda

Introduction

Practical Example

Summary

Bibliography

Research project “Cool Silicon”

1 out of 5 leading-edge clusters
within hightech strategy of the German Federal Government

- **Project term:**
  2009 – 2014
- **Budget:**
  135 million Euro

Members of Cool Silicon (Source: cool-silicon.de)

Aim: Substantially increase energy efficiency of ICs and information technology

Scope of a IT-subproject in “Cool Silicon”

Aims:

1) Performance improvement
2) Reducing energy consumption
3) New functionality for Business Intelligence

Methods:

- Investigation and testing of appropriate infrastructure and system architecture
- Definition and testing of appropriate modeling / algorithms
- Creation of prerequisites for efficient real-time aggregations and analysis (including ad hoc reporting)
- Creating prerequisites for efficient parallelization
- Reduction of complexity of nightly batch runs for data aggregation

Agenda

Introduction

Practical Example

Summary

Bibliography

Illustration of principles of real-time data analytics

Practical example: Computation of standard deviation s_N for N observations / items $x_1, x_2, x_3, \ldots, x_N$

- **Classical calculation:** (Cannot be used for real-time computation)

$$s_N = \Big(\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2\Big)^{1/2} \quad \text{with } \bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i.$$

  To calculate s_N red sum has to be recomputed from scratch.

- **Real-time calculation:**

$$s_N = \frac{1}{N}\Big(\Big|N\sum_{i=1}^{N} x_i^2 - \big(\sum_{i=1}^{N} x_i\big)^2\Big|\Big)^{1/2}.$$

  Obtained by multiplication and regrouping.
  If new item is added, then both sums can be updated.
  s_N will be calculated only on demand.

Agenda

Introduction

Practical Example

Summary

Bibliography

Classical vs. real-time aggregation

Classical aggregation (batch mode):
Calculation is started only when all the data involved becomes available (e.g., end of the day, etc.).

- Large amounts of data have to be calculated.
- Potentially complicated performance improvement algorithms are necessary.
- Race against time.

Real-time aggregation:
Data are aggregated as soon as data are known to the system.

- Small amounts of data are computed quasi continuously (e.g., in memory spread over the whole day).
- Smaller hardware is sufficient.
- Energy efficient.

Agenda

Introduction

Practical Example

Summary

Bibliography

Bibliography

Patent US 2017/0032016 A1
Zinner et al.: “Real-time Information Systems and Methodology Based on Continuous Homomorphic Processing in Linear Information Spaces”
`http://www.google.ch/patents/US20170032016`