Moderated Resource Elasticity for Stream Processing Applications

Michael Borkowski, Christoph Hochreiner, Stefan Schulte

Distributed Systems Group
TU Wien
Vienna, Austria

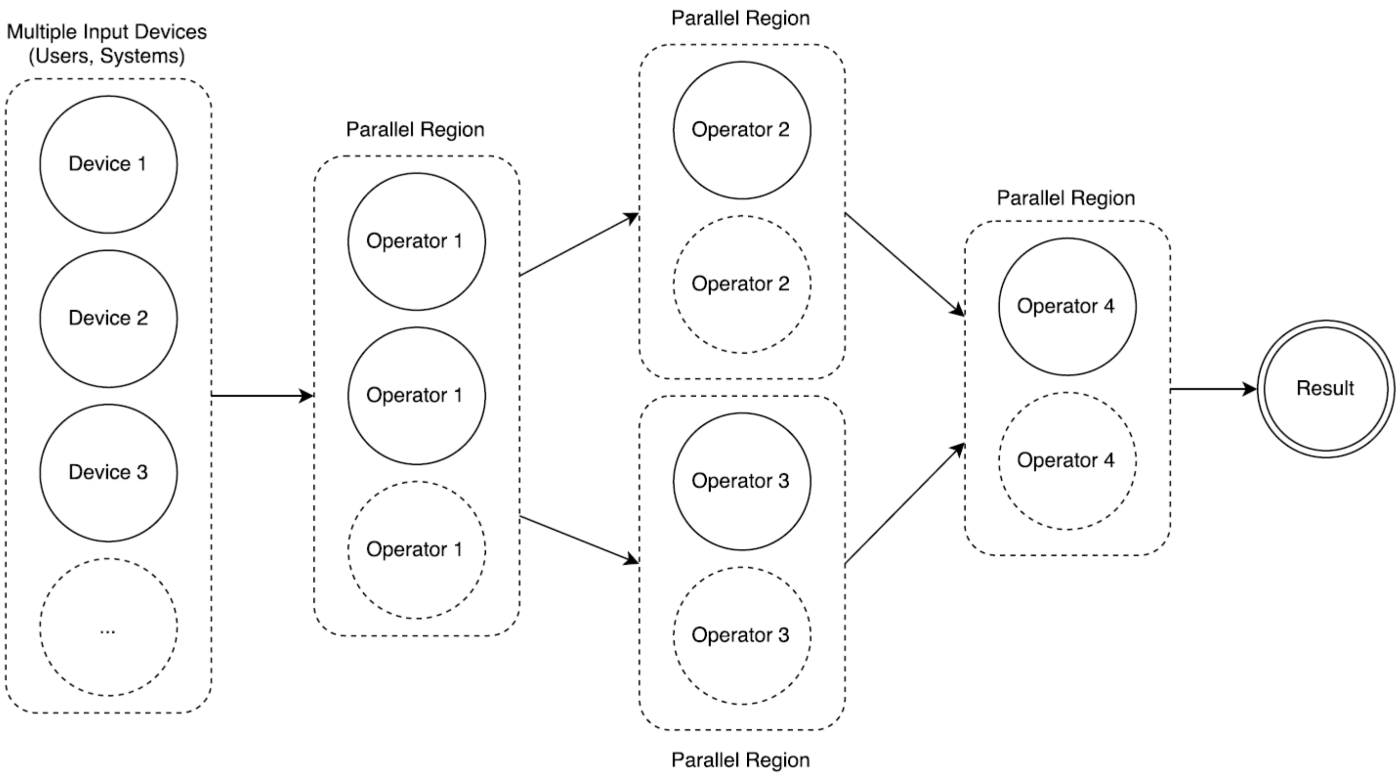
August 29th 2017

International Workshop on Autonomic Solutions for Parallel and Distributed Data Stream Processing (Auto-DaSP 2017)
at 23rd International European Conference on Parallel and Distributed Computing (Euro-Par 2017)
Santiago de Compostela, Spain

Distributed Systems Group

Introduction

Motivation Scenario



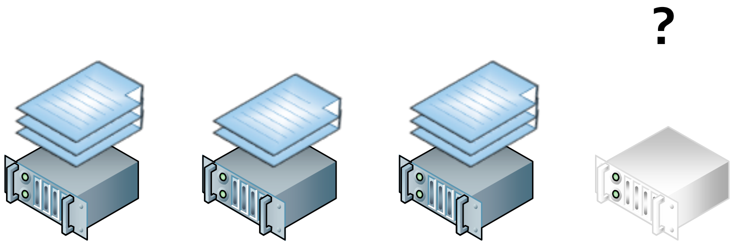
Roland Doppler: Scaling Algorithms for Distributed Stream Processing Applications

Scaling

"Do I need to activate more workers to process X, Y and Z?"

"Are workers running unnecessarily that can be shut down?"

- Insufficient workers: Performance drops, SLA violations, increased penalties
- Excessive workers: Unnecessary resources running, increased cost

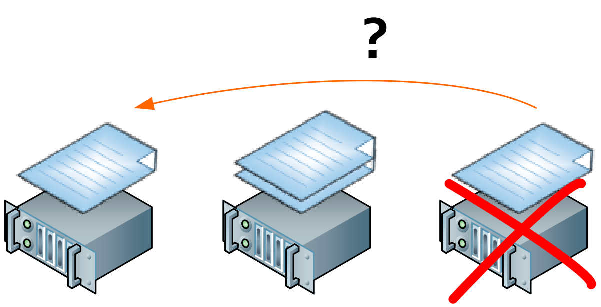


Scaling

"*Do I need to activate more workers to process X, Y and Z?*"

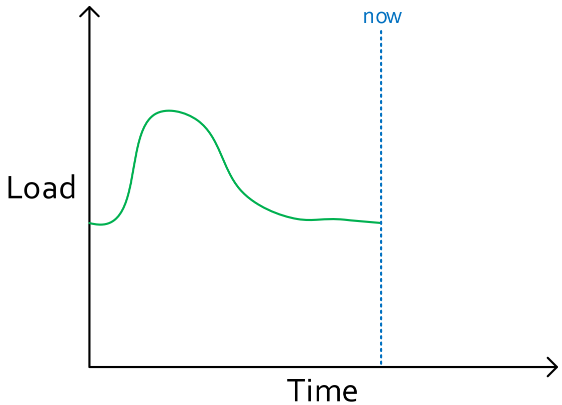
"*Are workers running unnecessarily that can be shut down?*"

- Insufficient workers: Performance drops, SLA violations, increased penalties
- Excessive workers: Unnecessary resources running, increased cost



Reactive vs. Proactive Approaches

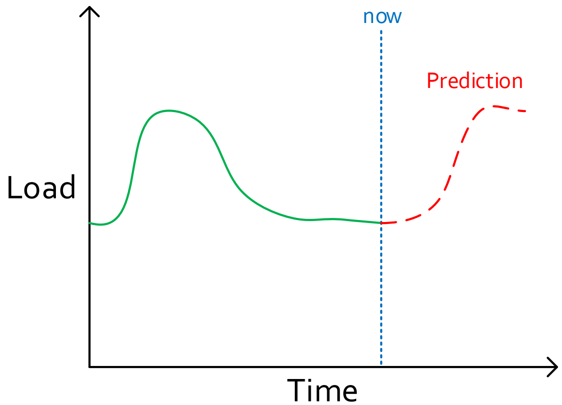
Reactive



Simpler design

No prediction overhead

Proactive



Potential for in-time decisions

Necessary reaction can be estimated

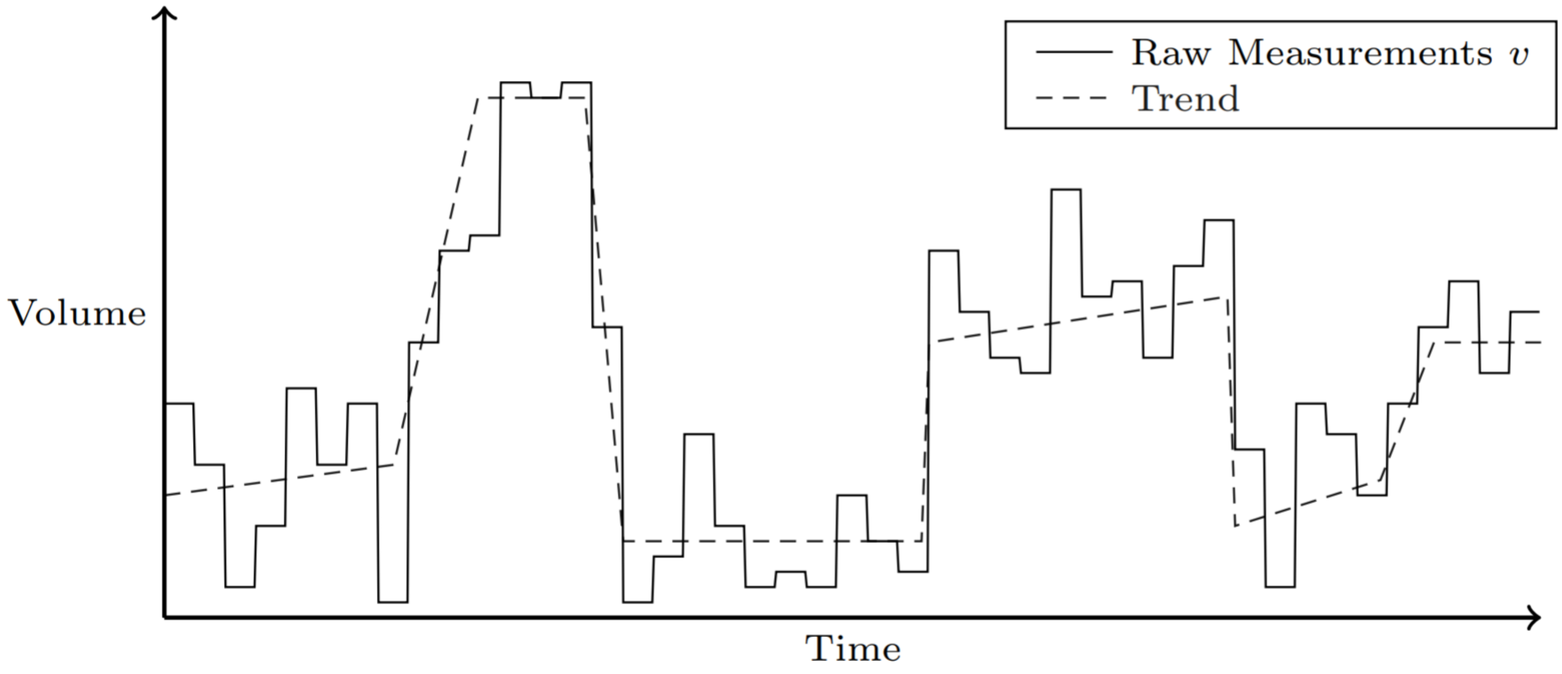
Approach

Background

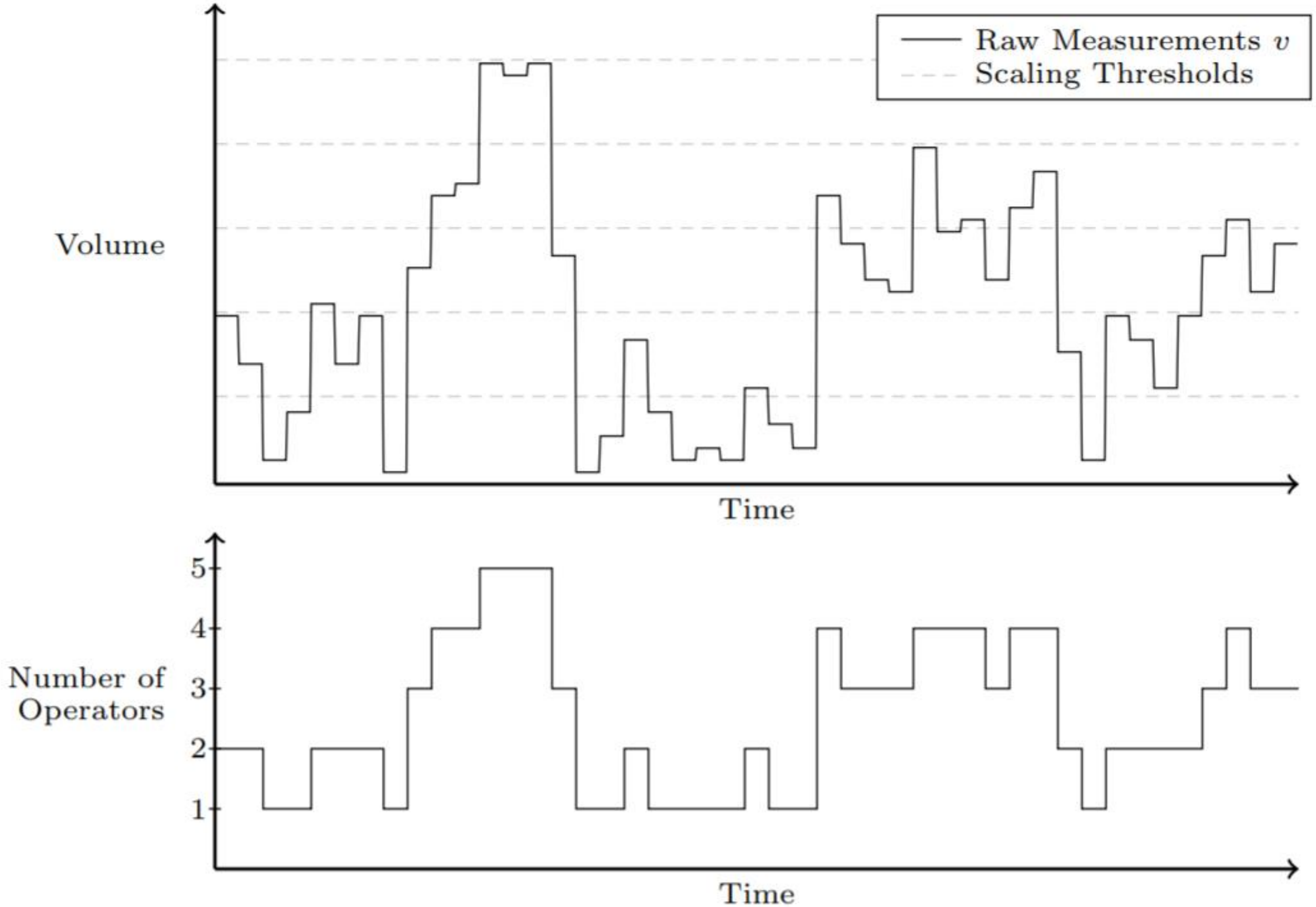
Overall Goal: Prediction of input volume

Challenge: Noise in data

Background

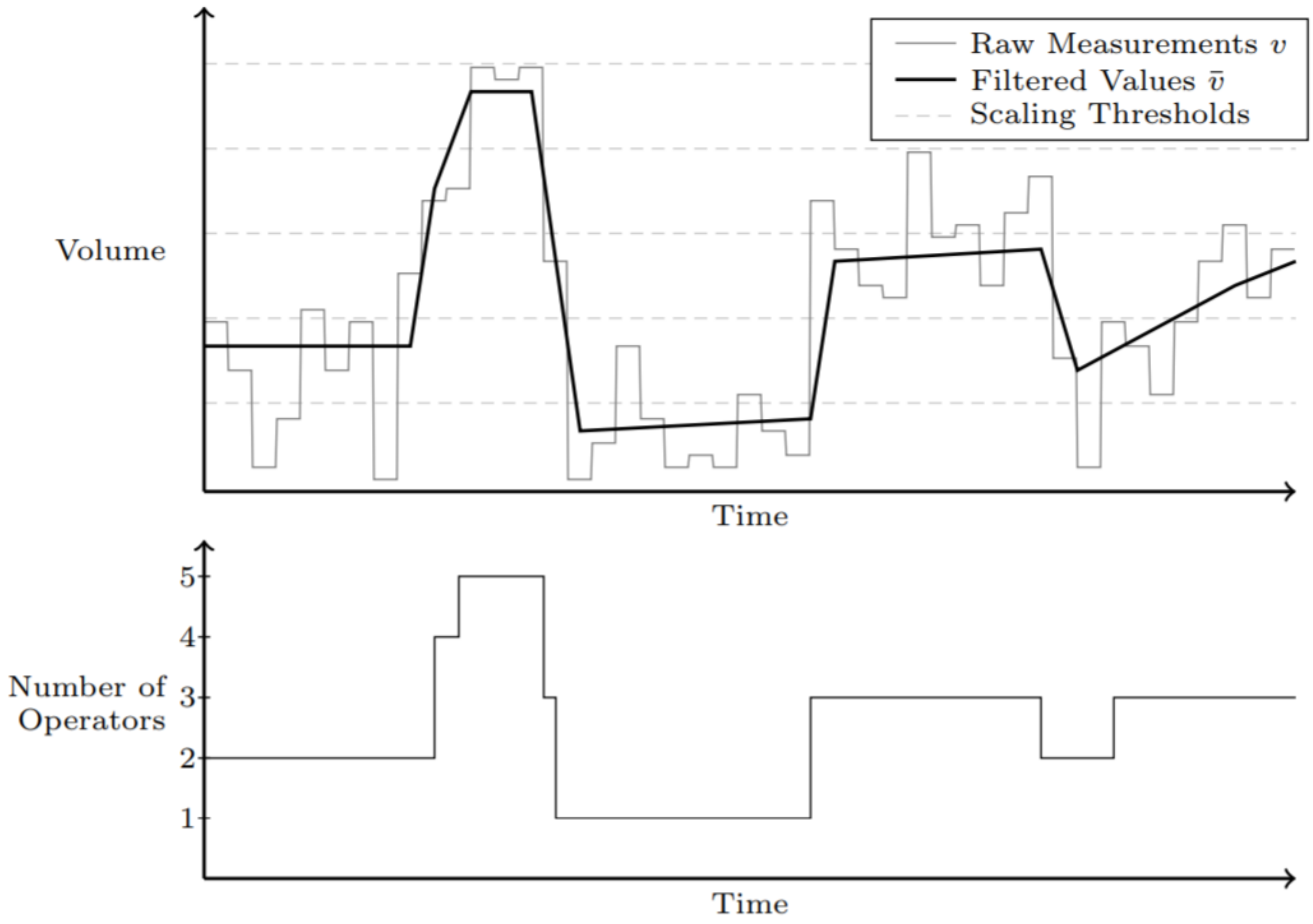


Background



Raw Measurements v
Scaling Thresholds
Volume
Time
Number of Operators
5
4
3
2
1
Time

Background



Raw Measurements v
Filtered Values $\bar{v}$
Scaling Thresholds
Volume
Time
5
4
3
2
1
Number of Operators
Time

Approach

Solution within current paper:

- Total Variation Denoising [1]
- Extended Kalman Filter [2]

[1] Rudin, Leonid I., Stanley Osher, and Emad Fatemi. "Nonlinear total variation based noise removal algorithms." Physica D: Nonlinear Phenomena 60.1-4 (1992): 259-268.

[2] Kalman, Rudolph E., and Richard S. Bucy. "New results in linear filtering and prediction theory." Journal of basic engineering 83.1 (1961): 95-108.

Approach: TVD

Two constraints:

- Preserve mean of raw signal

$$\int_{\Omega} u = \int_{\Omega} u_0$$

- Define deviation of noise

$$\int_{\Omega} (u - u_0)^2 = \sigma^2$$

Approach: TVD

Two constraints:

- Preserve mean of raw signal

$$\int_{\Omega} u = \int_{\Omega} \boxed{u_0}$$

u_0: raw signal

- Define deviation of noise

$$\int_{\Omega} (u - \boxed{u_0})^2 = \sigma^2$$

Approach: TVD

Two constraints:

- Preserve mean of raw signal

$$\int_{\Omega} u = \int_{\Omega} u_0$$

u: filtered signal

- Define deviation of noise

$$\int_{\Omega} (u - u_0)^2 = \sigma^2$$

Approach: TVD

Two constraints:

- Preserve mean of raw signal

$$\int_{\Omega} u = \int_{\Omega} u_0$$

- Define deviation of noise

$$\int_{\Omega} (u - u_0)^2 = \sigma^2$$

Approach: TVD

Minimization of variation:

$$\text{minimize} \int_{\Omega} \left(u_{xx} + u_{yy}\right)^2$$

Approach: TVD

SNR = 1.0

Approach: EKF

Initial estimate

Predict

1. Project the state
2. Project the error covariance

Measure

1. Compute Kalman gain
2. Update estimate with actual measurement

Evaluation

Evaluation: Setup

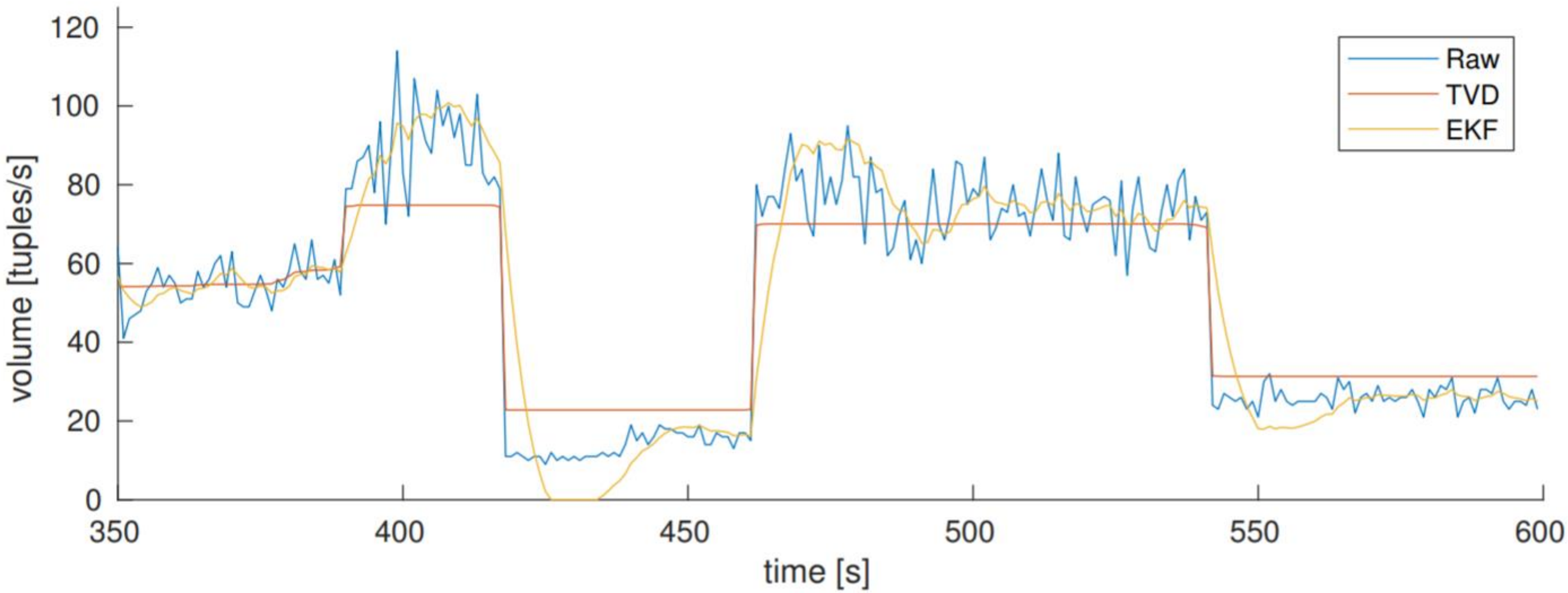
Simulation of a stream processing engine

Values for data volume presented in [3]:

- Data rate: 200–500 tuples per second
- Operator capacity: 50 tuples per second
- SNR: 0.5, 0.8, 1.0, 5.0, 10.00
- Simulation duration: 1500 seconds
- Measurement interval: 1 second

[3] Xu, Jielong, et al. "T-storm: Traffic-aware online scheduling in storm.", 34th International Conference on Distributed Computing Systems (ICDCS). IEEE, 2014.

Evaluation



120
100
80
60
40
20
0
volume [tuples/s]
350
400
450
500
550
600
time [s]
Raw
TVD
EKF

Evaluation: Metrics

Baseline: Linear Smoothing

Comparison to both TVD and EKF

Measured Variables:

- $\mathbf{s}^{+}$ and $\mathbf{s}^{-}$: scaling operations
- $\mathbf{p}^{+}$ and $\mathbf{p}^{-}$: provisioning state (over- and under-provisioning)

Evaluation: Results

	Scaling Operations		Provisioning Time [s]	
Filter	**s^+**	**s^-**	**p^+**	**p^-**
Linear Smoothing (Baseline)	130	130	186	<u>220</u>
Total Variation Denoising	<u>8</u> *-122*	<u>7</u> *-123*	<u>154</u> *-32*	224 *+4*
Extended Kalman Filter	79 *-51*	78 *-52*	319 *+133*	224 *+4*

SNR = 1.0
duration = 1500 s

Conclusions & Future Work

Conclusions

Contributions

- Application of TVD and EKF within stream processing
- Simulation-based evaluation
- With current setup: TVD is promising

Future Work

- Work on EKF models to improve EKF performance
- Employ a feedback loop
- Usage of gained smoothed data for predictions
- Testbed evaluation using VISP [4]

[4] Hochreiner, Christoph, et al. "VISP: An Ecosystem for Elastic Data Stream Processing for the Internet of Things.", 20th International Enterprise Distributed Object Computing Conference (EDOC). IEEE, 2016.

?

Michael Borkowski

Distributed Systems Group · TU Wien

m.borkowski @ infosys.tuwien.ac.at