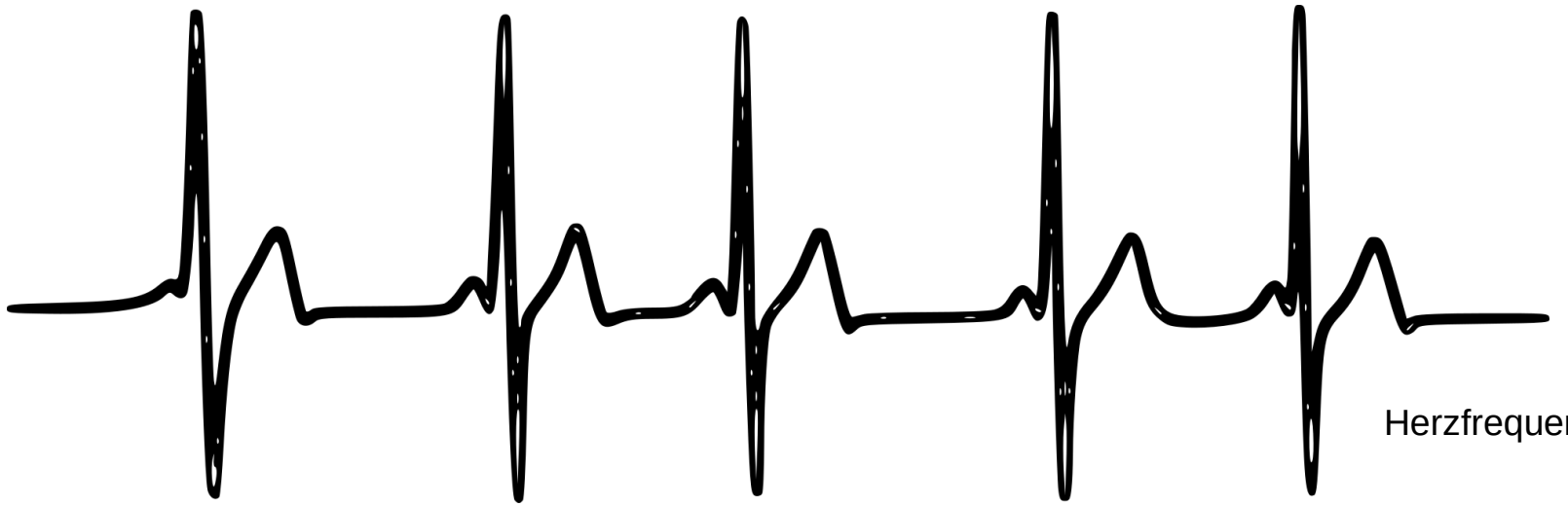
Anomaly Detection

für unternehmenskritische Applikationen in heterogenen IT-Umgebungen

Susanne Greiner

Einführung

Herzfrequenz



Herzfrequenz

Abhängig von

- Pathologie
- Sport
- Atmung
- Pharmaka
- Temperatur
- Dehydration
- Druck
- Etc.

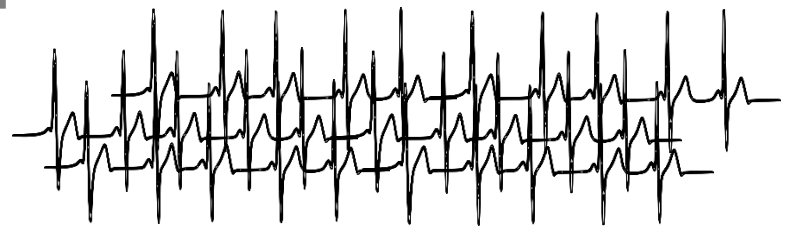
Monitoring & Alarme

Personenspezifische Daten

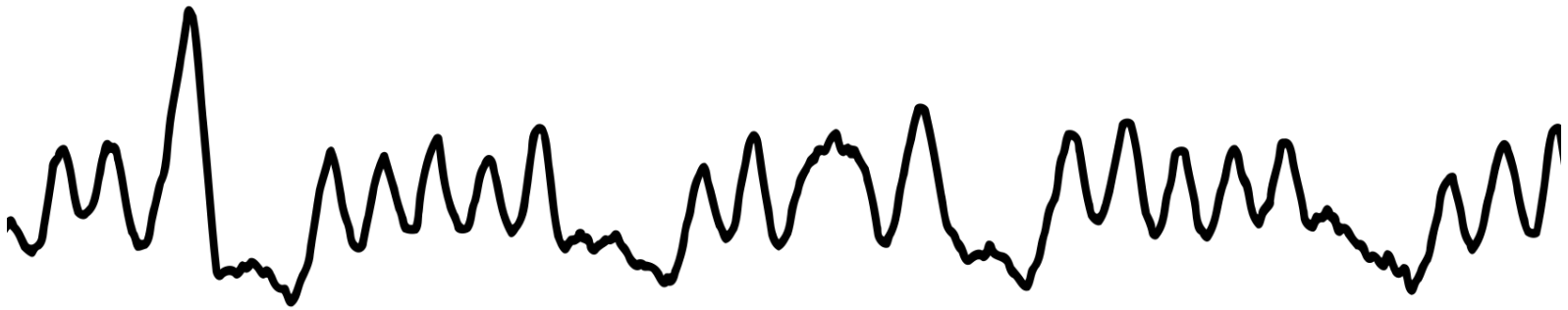
Erwartung

- Zeitreihe
- Alarm Schwellwerte

Bevölkerungsdaten



Performance Monitoring



Prozent Prozessorzeit

Abhängig von

- Batch requests
- Transactions
- Memory
- SAN
- Network
- Side Processes
- Etc.

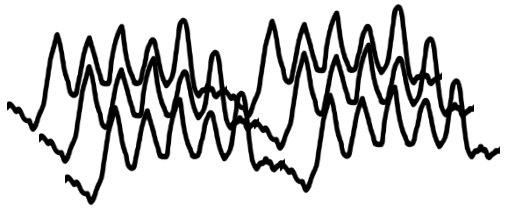
Monitoring & Alarme

Spezifische Maschinen Daten

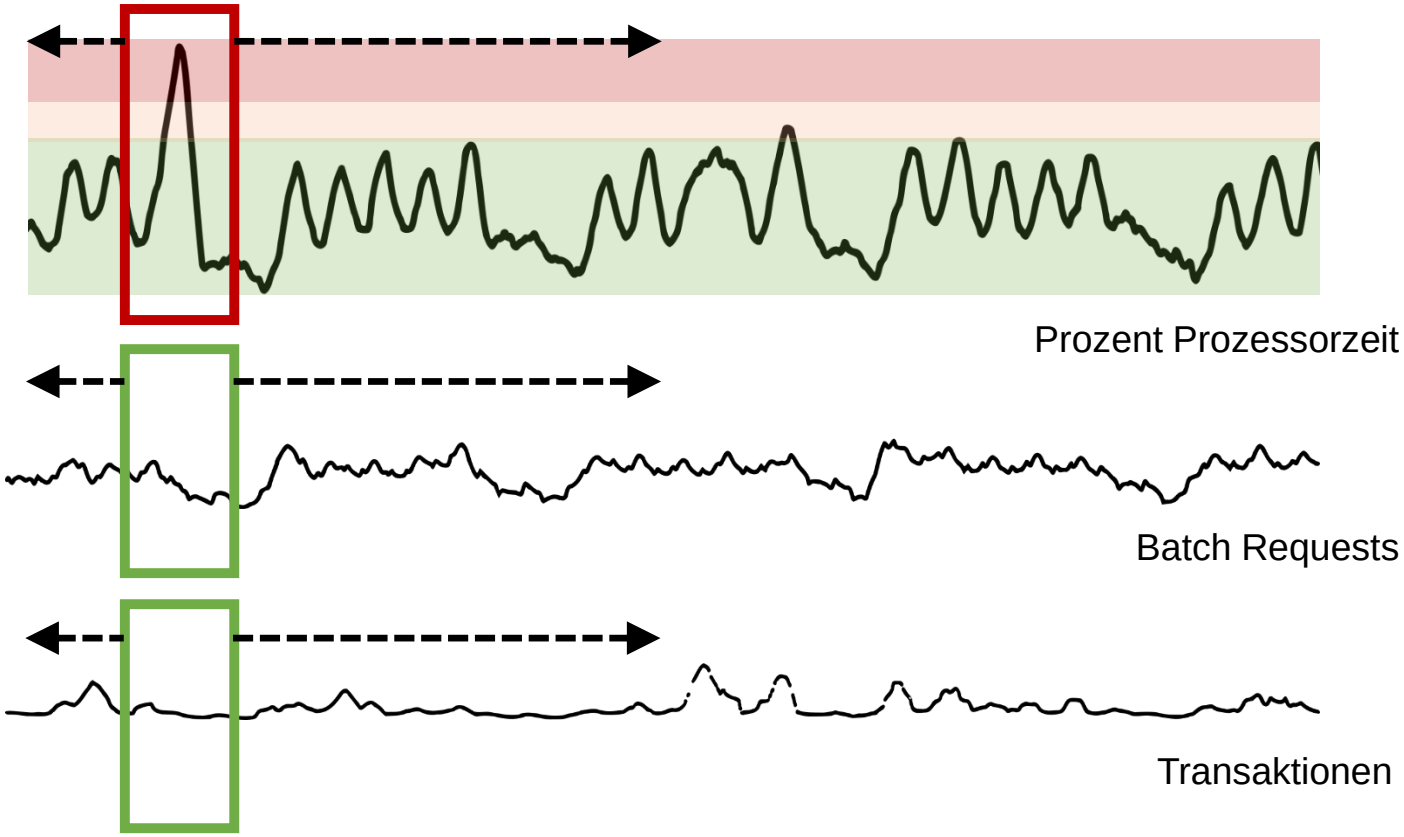
Erwartung

- Zeitreihe
- Alarm Schwellwerte

Erfahrung,
Daten ähnlicher Machines/ Settings



ABER: State of the Art Performance Monitoring

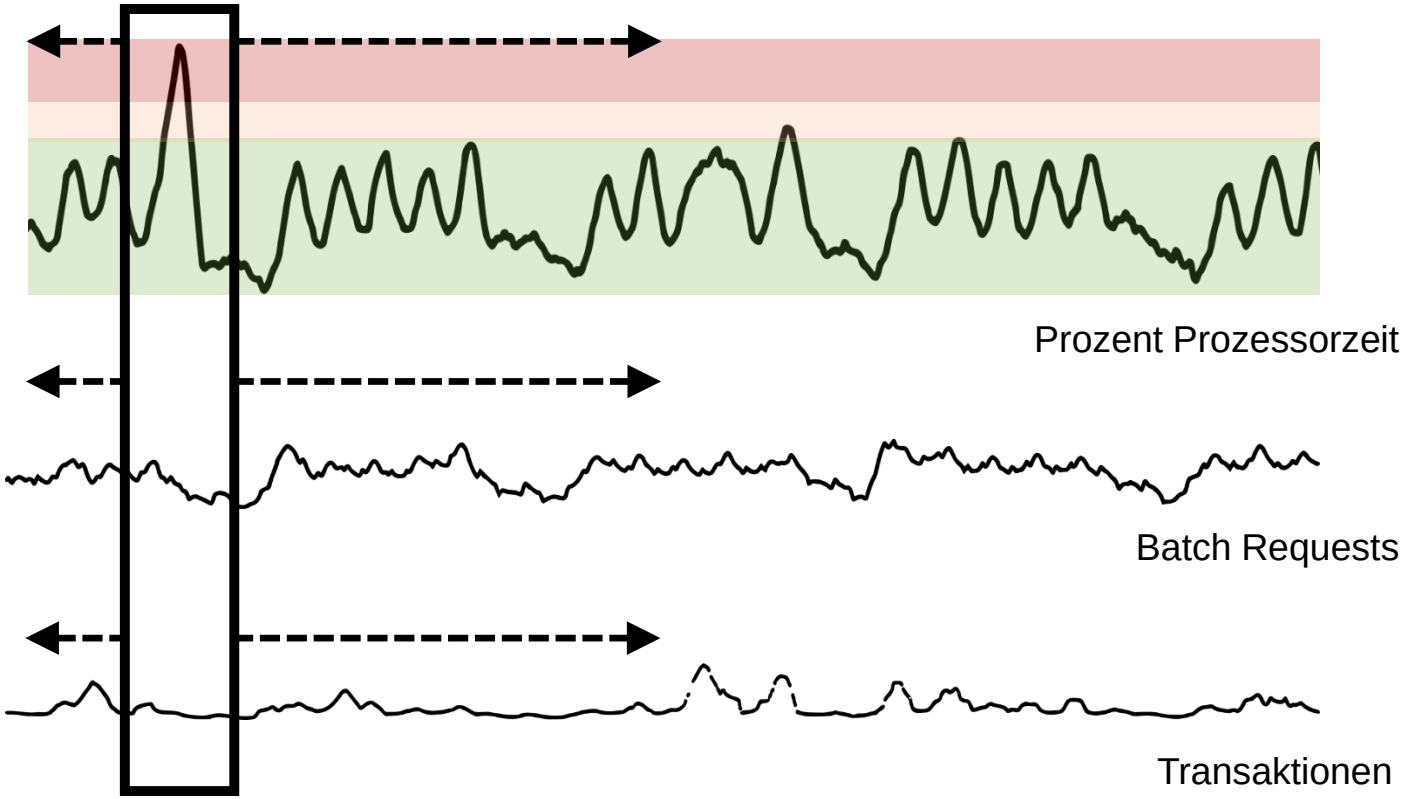


- Jede Zeitreihe wird **einzeln** analysiert
- Schwellwerte werden via **Baselining** berechnet
- Alarme verschiedener Zeitreihen werden zu einem **globalen Alarm** kombiniert
- Abhängigkeiten zwischen den Zeitreihen werden **ignoriert**
- Form der Zeitreihe wird **ignoriert**

Motivation

- Unterschiedliche Datenquellen
- Unterschiedliche Aufzeichnungsgenauigkeiten
- Evolution der Netzwerkkomplexität: Common Practice **war** genug

Next Level Performance Monitoring



Etc.

- Alle Zeitreihen werden **zusammen** analysiert
- Schwellwerte werden via **Baselining und Anomaly Detection** errechnet
- **Risk Berechnung** zusätzlich zu **globalen und spezifischen Alarmen**
- **Abhängigkeiten** zwischen Zeitreihen werden verwendet um die Alarmqualität und Risks zu verbessern
- Form der Zeitreihe wird berücksichtigt

Motivation

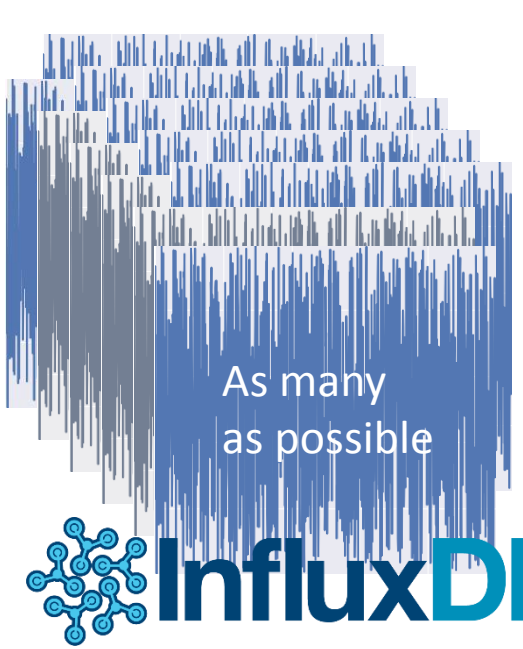
- Gemeinsame Datenquelle
- Grafana & InfluxDB
- Heute braucht es mehr als Common Practice
- **Proactivity**

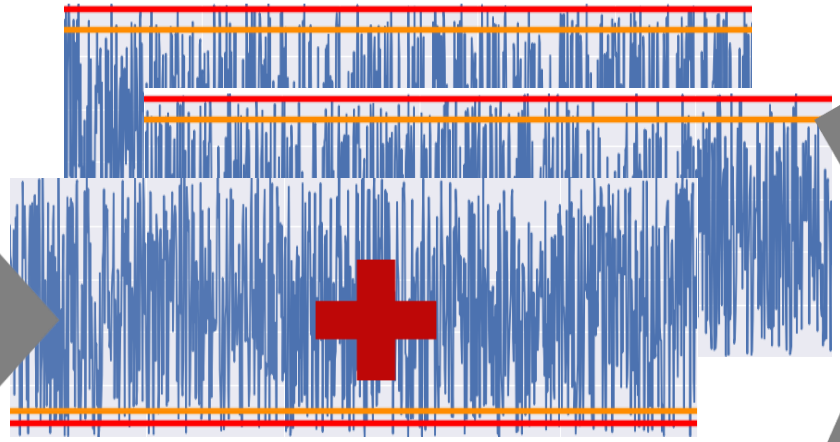
Next Level Performance Monitoring

Metriken

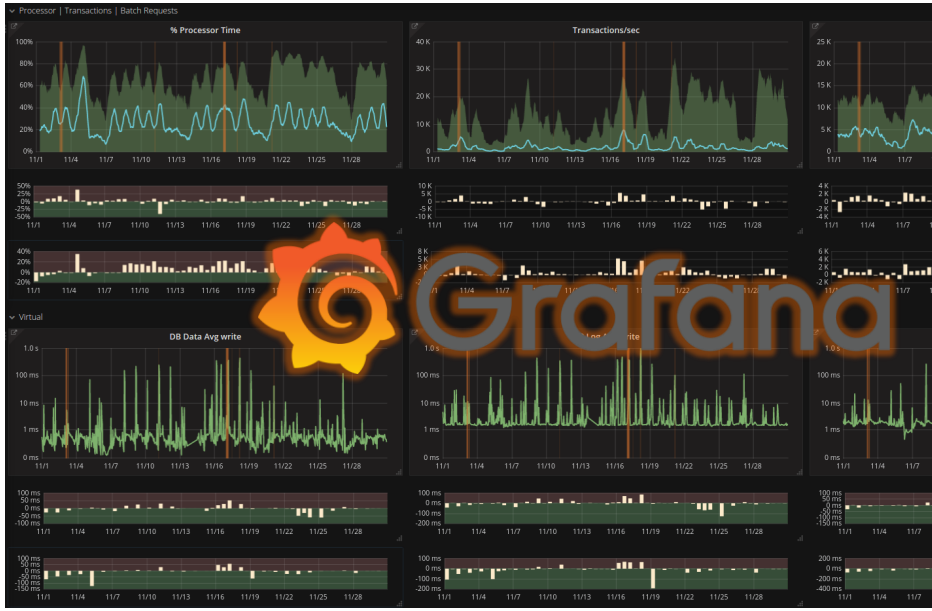
Baselining

Visualisierung





MACHINE LEARNING
ADVANCED STATISTICS

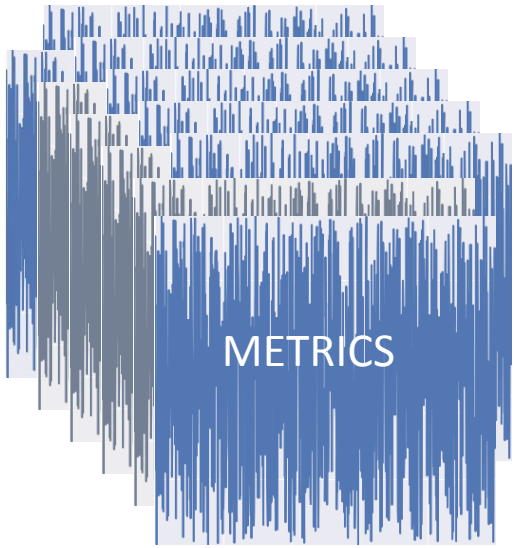


Risk

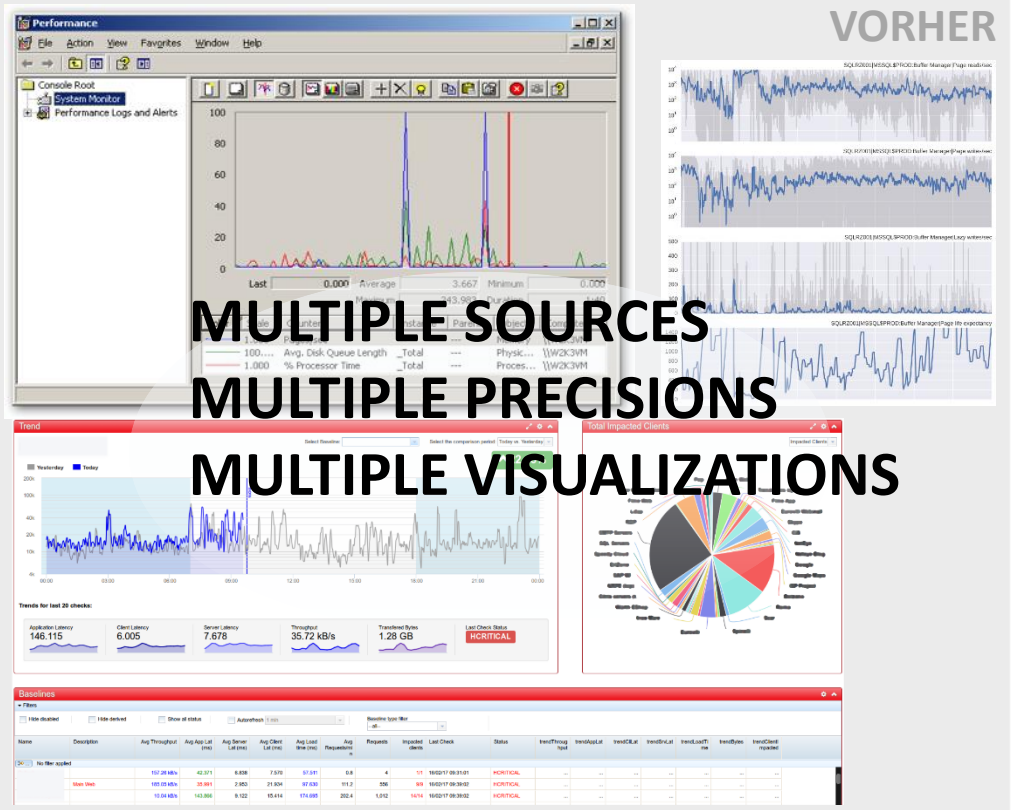
Alerts

Details

Eine gemeinsame Datenquelle

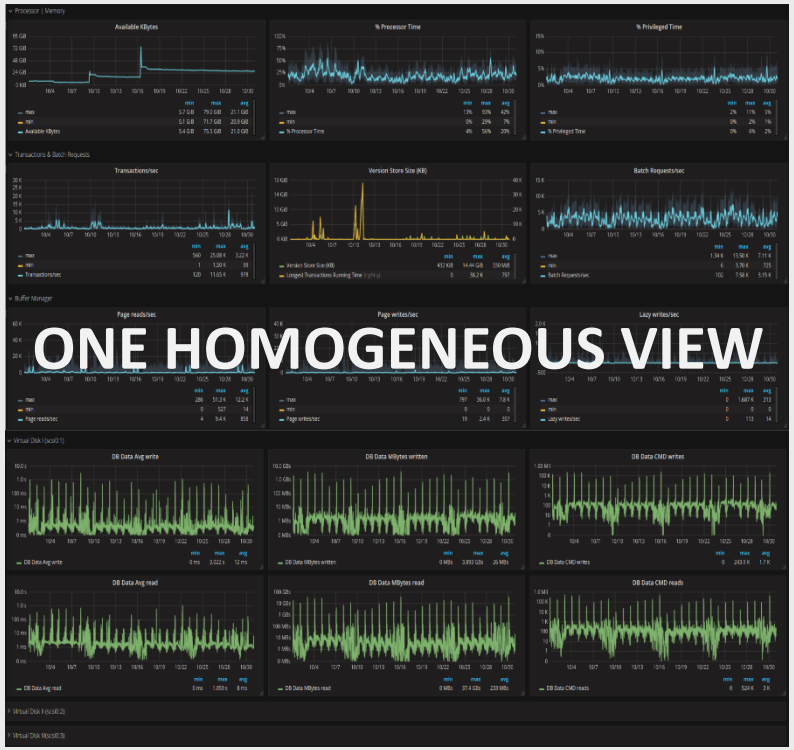


VORHER



MULTIPLE SOURCES
MULTIPLE PRECISIONS
MULTIPLE VISUALIZATIONS

NACHHER



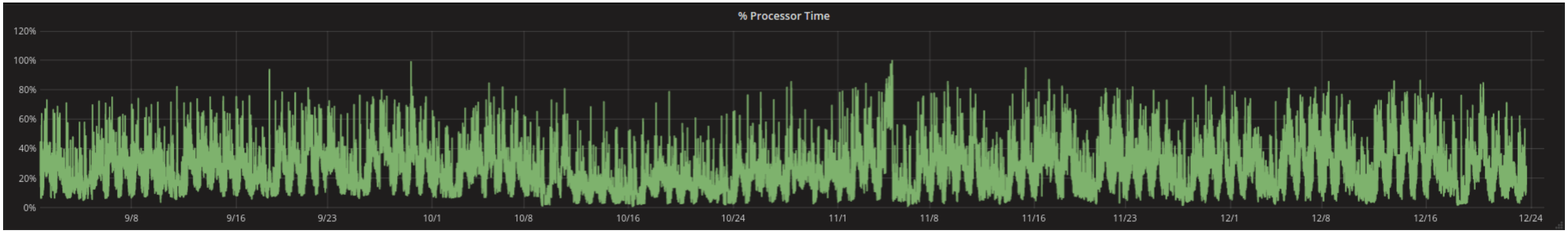
ONE HOMOGENEOUS VIEW

InfluxDB

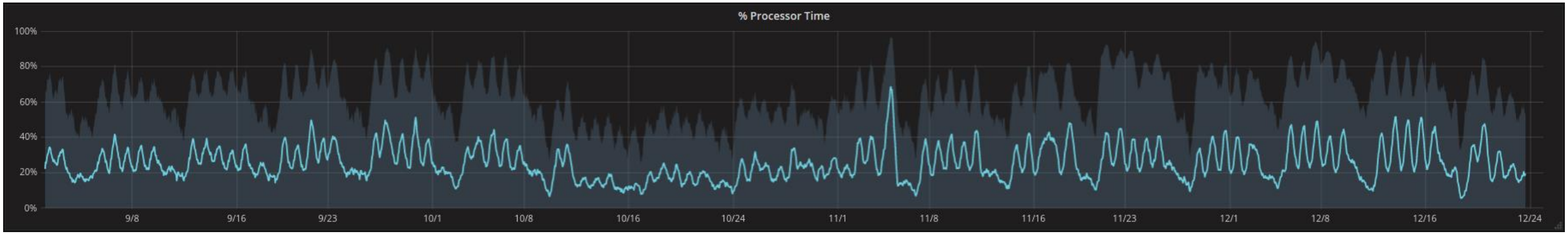
- Kommunikation mit nur einer Datenquelle
- Verstehen nur eines Outputtyps
- Vergleich homogener Daten

Einfache Trenderkennung

VORHER



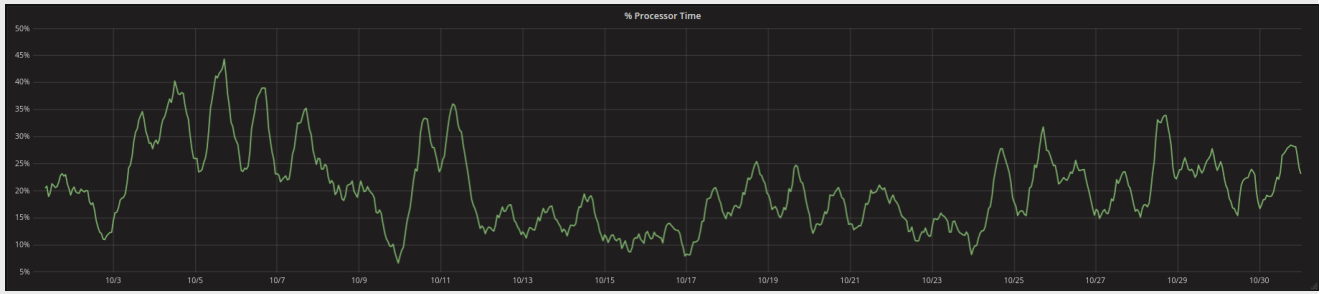
NACHHER



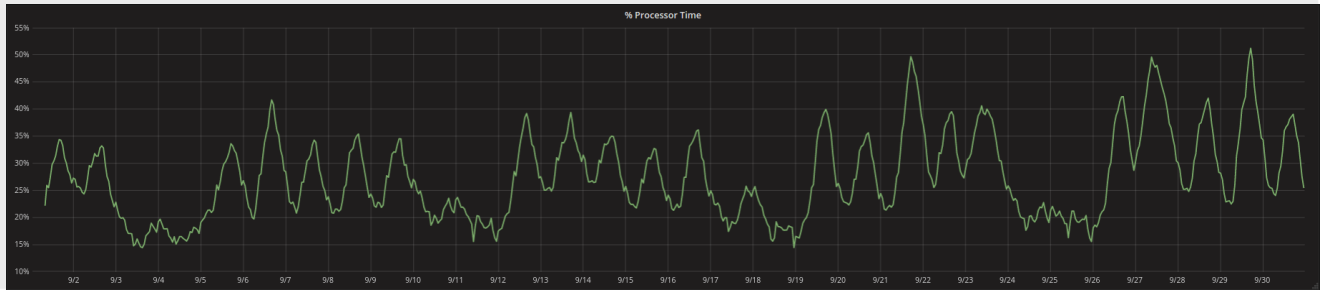
- Bereiche statt (zu) vieler Punkte
- Geglättetes Signal

Historische Daten

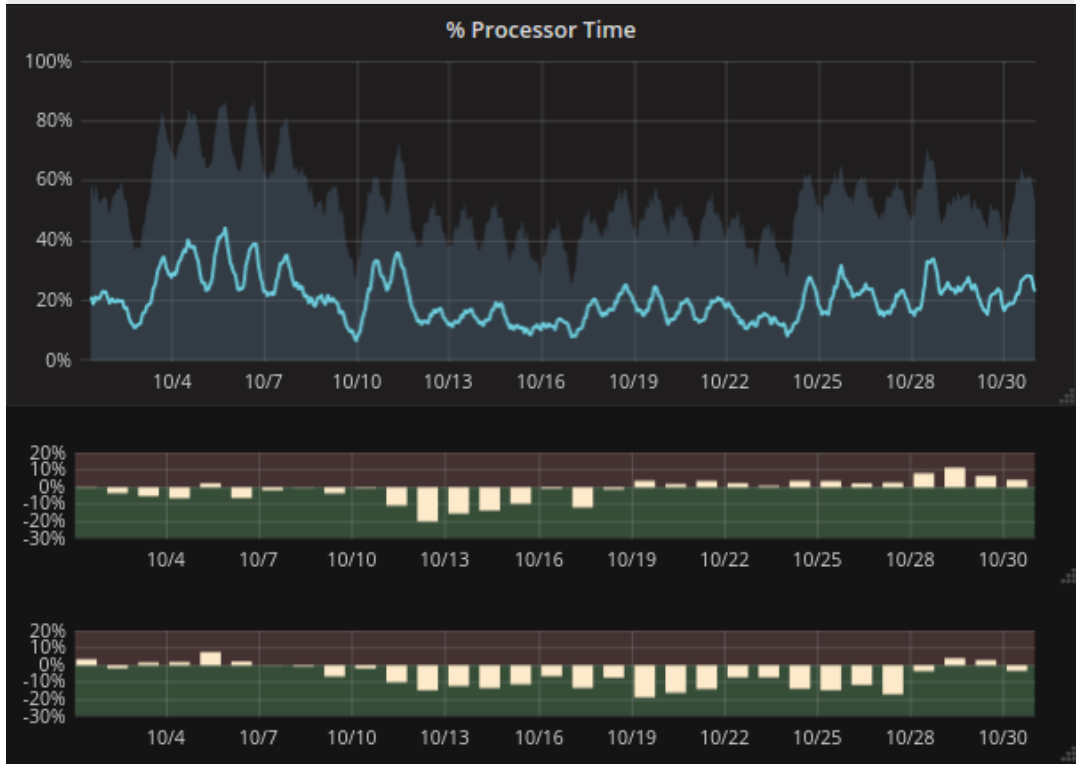
VORHER



- Vergleich von Zeitintervallen schwierig
- Quantifizierung von Effekten kaum möglich



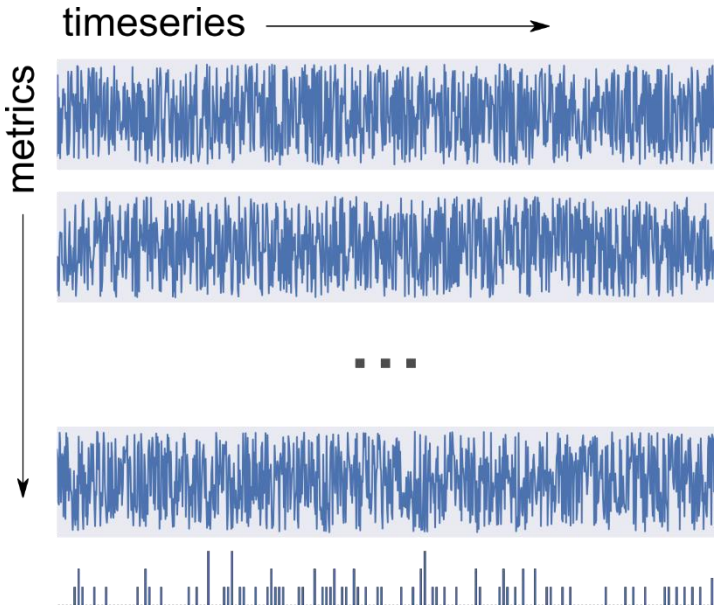
NACHHER



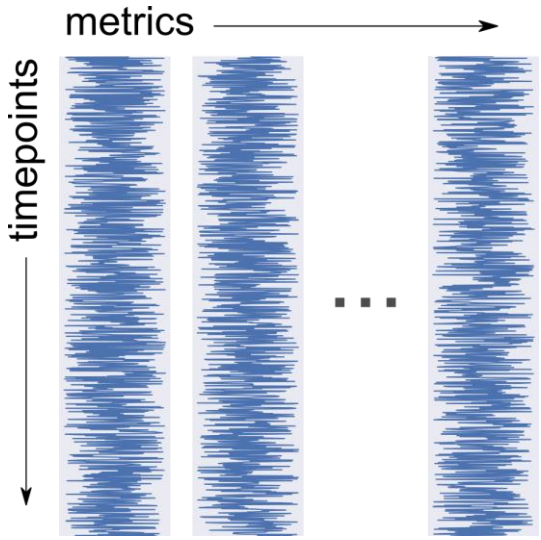
- Historische Daten griffbereit
- Differenzenvisualisierung
  => Einfache Effektquantifizierung

Risk als Ergänzung der Standard Alerts: Berechnung

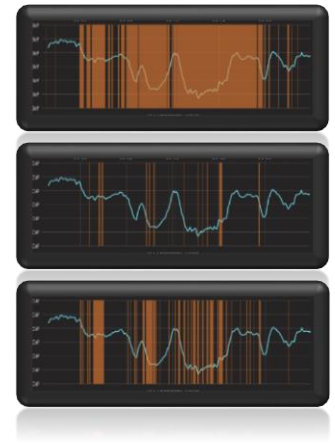
DATASET



FEATURE SPACE
jeder Zeitpunkt: Zustand, der eine *Standard* | *nicht Standard* Situation charakterisiert



MVML

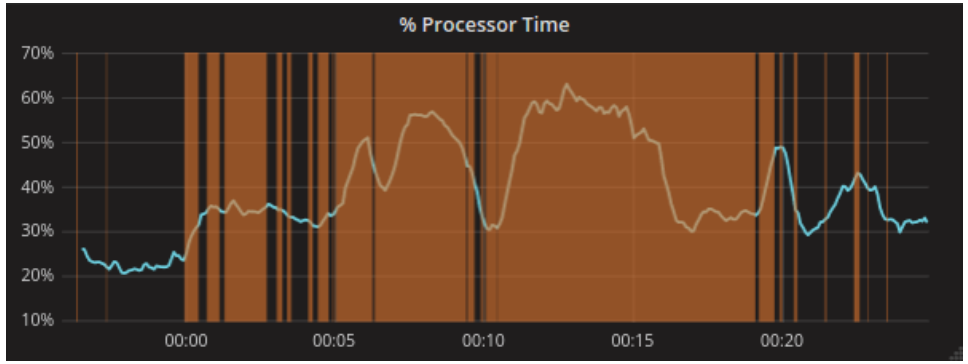


OUTPUT
RISK:
Anders als der 'Standard'

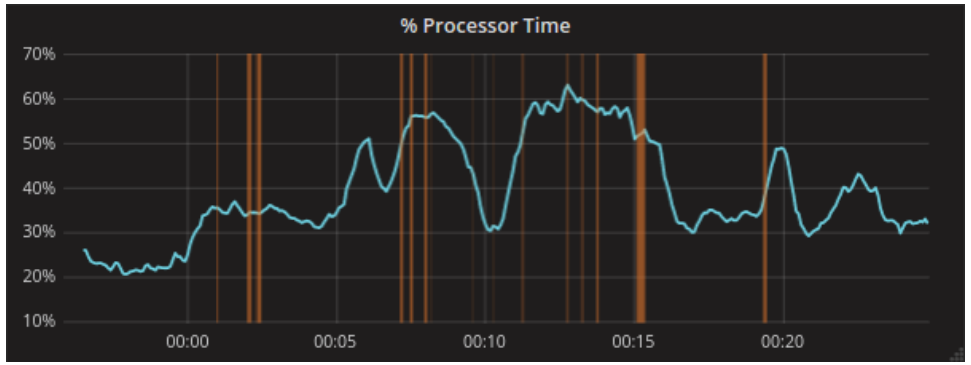
Standard: jede Zeitreihe wird einzeln analysiert und mit ihrem eigenen Standardverhalten verglichen

Multivariate analysis: jeder Zeitpunkt wird analysiert, dabei werden neben den einzelnen Werten jeder Metrik ihre Gesamtkonstellation betrachtet

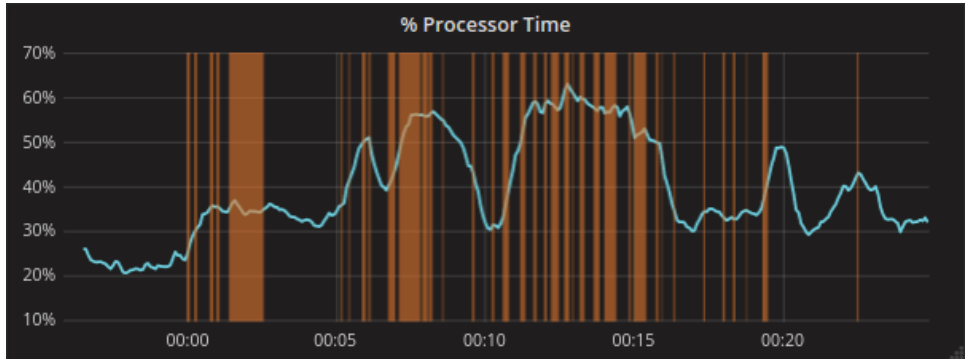
Risk als Ergänzung der Standard Alerts



Globales Risk:
Zeitpunkte mit Werten, die bezogen auf die Metriken, die herangenommen warden, merklich anders sind wenn man sie mit historischen Daten vergleicht

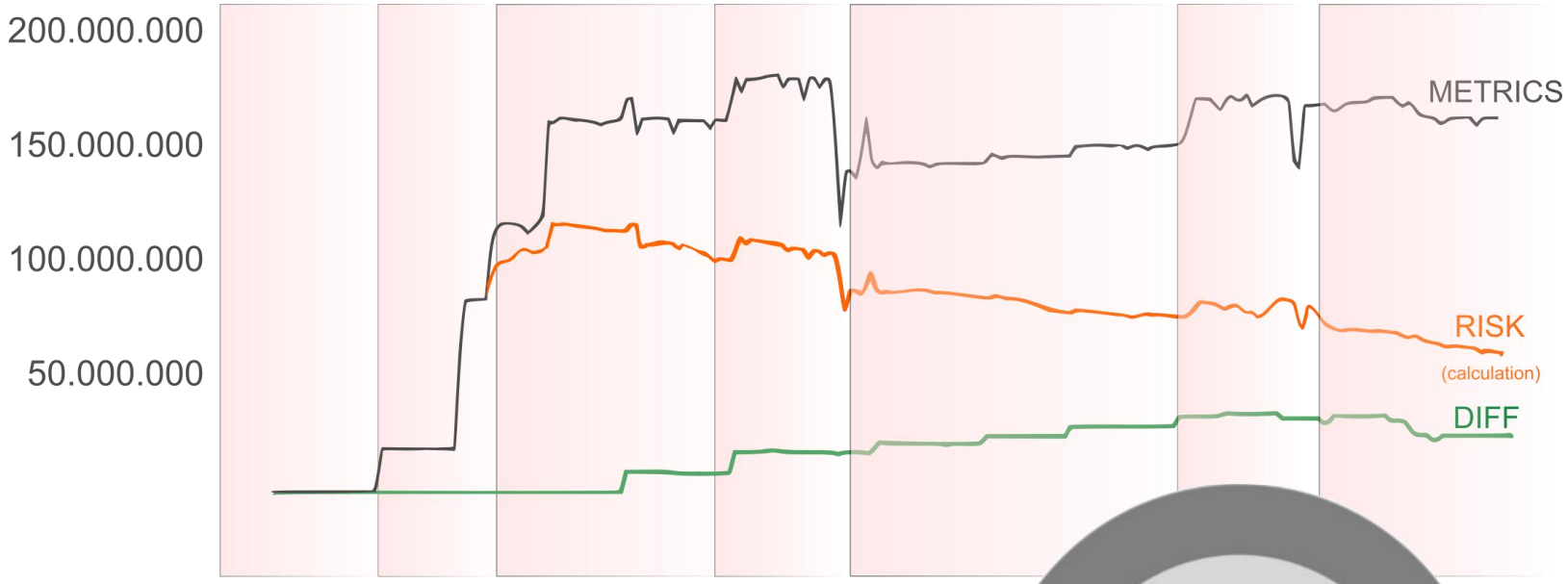


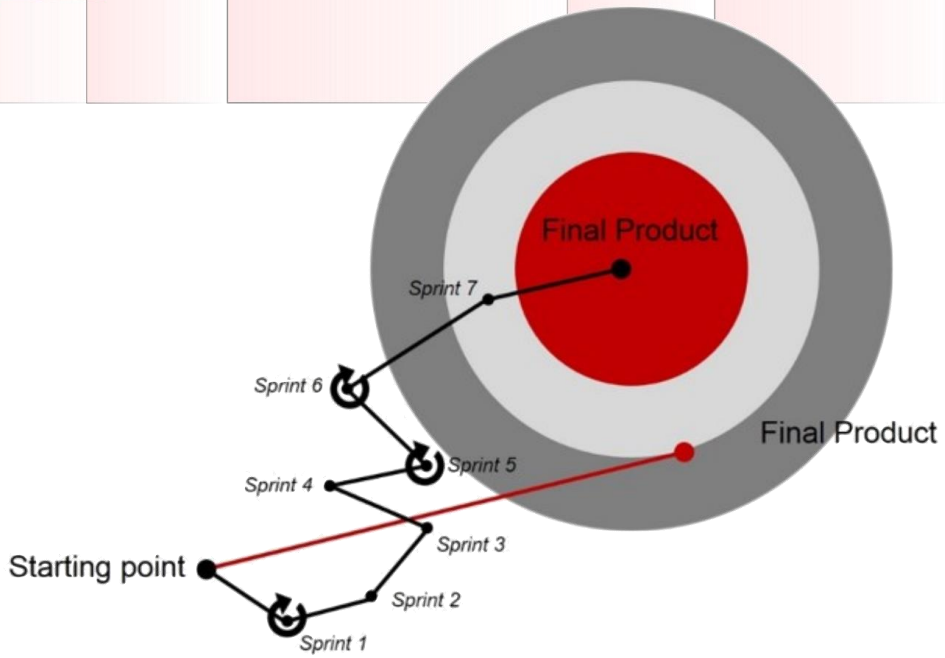
Spezifisches Risk:
Zeitpunkte deren Wertekonstellation nicht dem Standard entspricht verglichen mit historischen Daten des selben Tages



Duration Risk:
Zeitpunkte deren Wertekonstellation nicht dem Standard entspricht verglichen mit historischen Daten des selben Tages, die zusätzlich in der Nähe weiterer nicht Standard Datenpunkte liegen

Agile Implementation



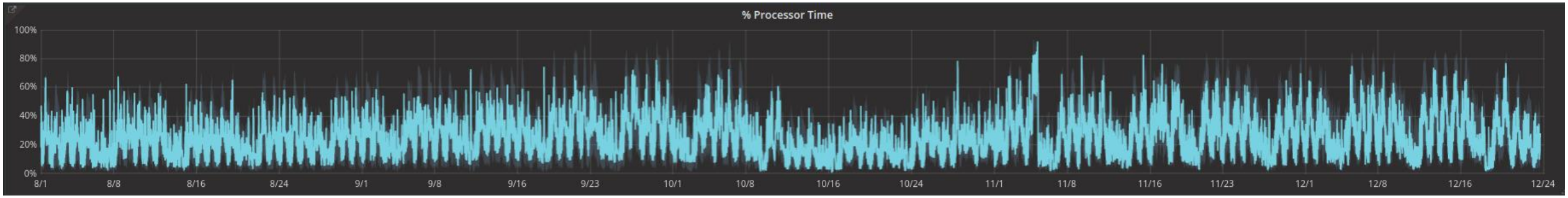


Lösung wird dank agiler Implementierung sehr kundenspezifisch

Use Cases

„Ist Zustand“ und Dynamik meines Systems?

VORHER

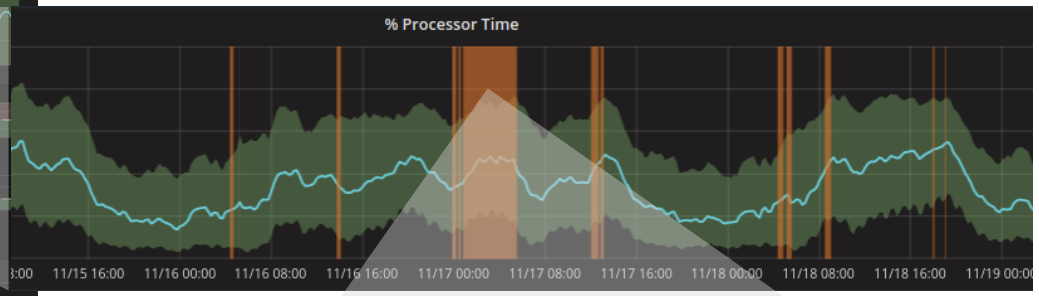


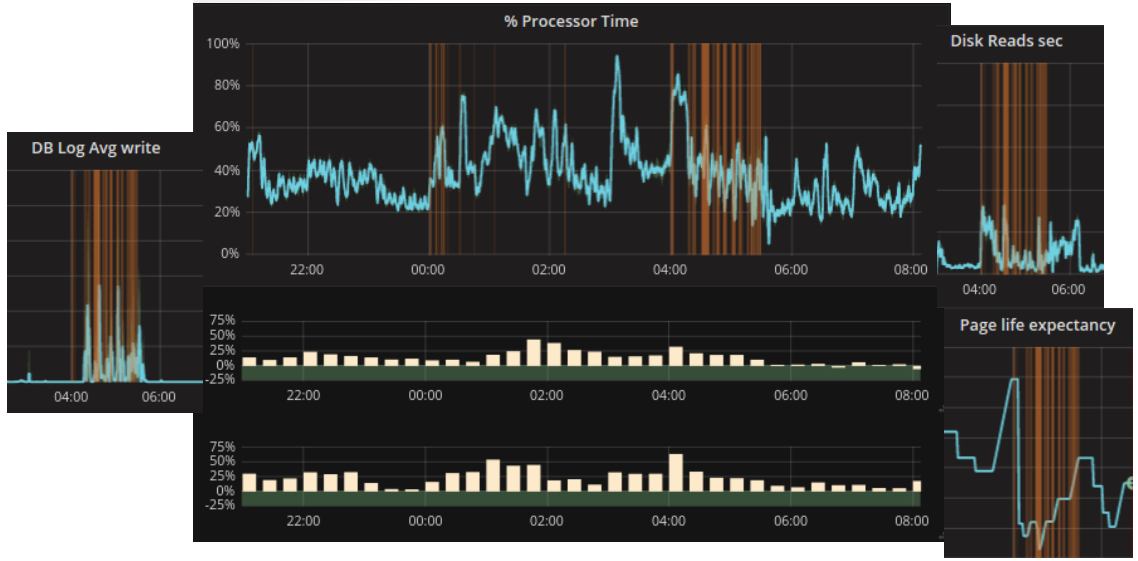
Daten können schneller und einfacher interpretiert werden

Auffälligkeiten können einfach kommuniziert und geteilt werden

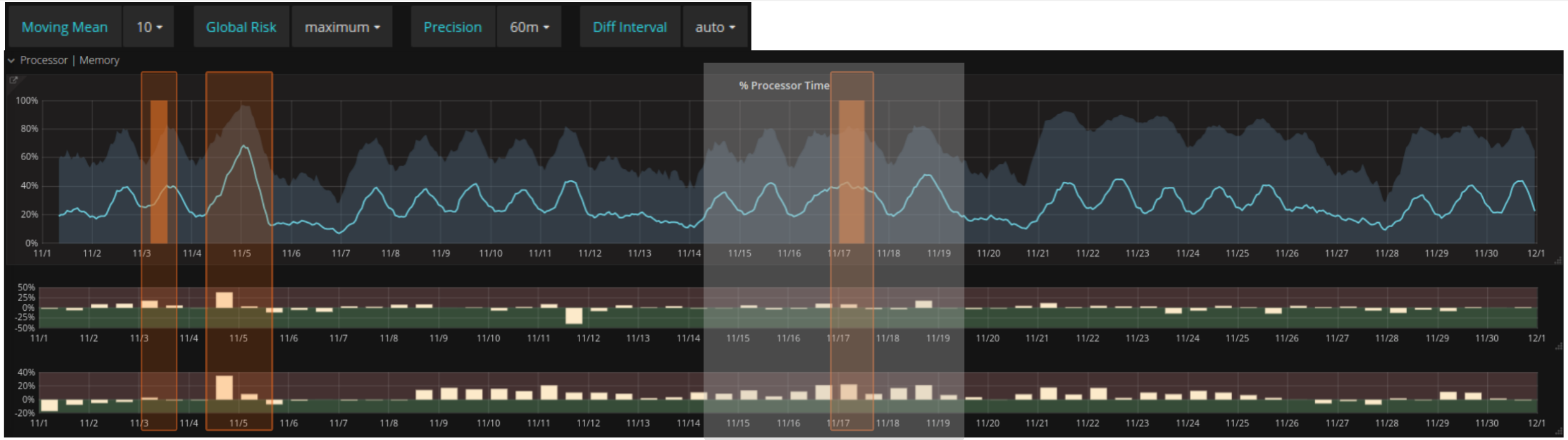
NACHHER

Troubleshooting





Beispiel I: Large Scale Fehleranalyse

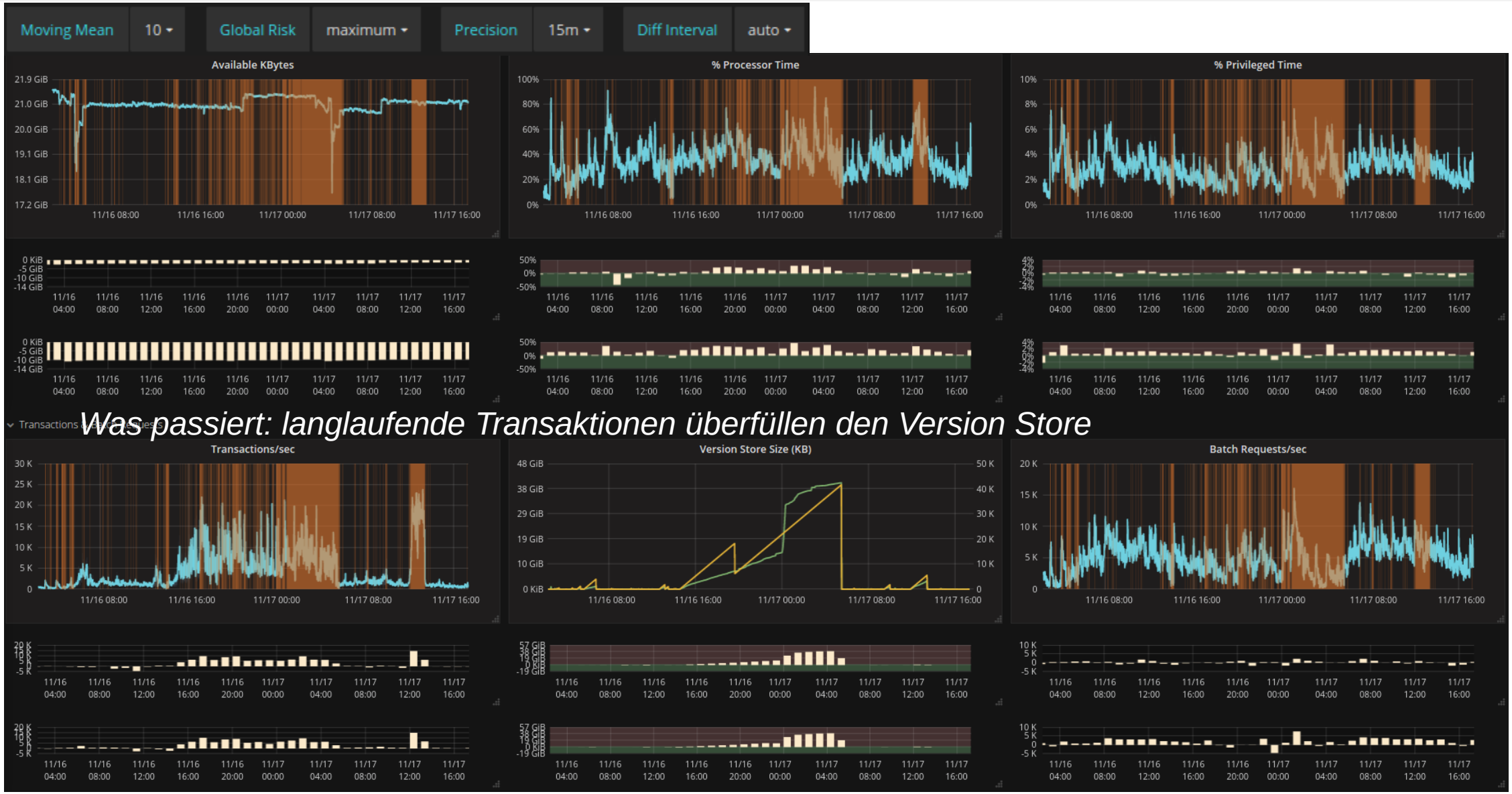


Ein ganzer Monat Daten, wo macht Fehleranalyse am meisten Sinn ?

Beispiel I:


Moving Mean
10
Global Risk
maximum
Precision
15m
Diff Interval
auto
Processor | Memory
% Processor Time
Transactions/sec
Batch Requests/sec
Erhöhte Transaktionen
So weder in der Vorwoche
noch im Vormonat vorhanden

Beispiel I:



Was passiert: langlaufende Transaktionen überfüllen den Version Store

Beispiel I:

Scheinen nicht direkt aus den Monitoring Metriken erklärbar

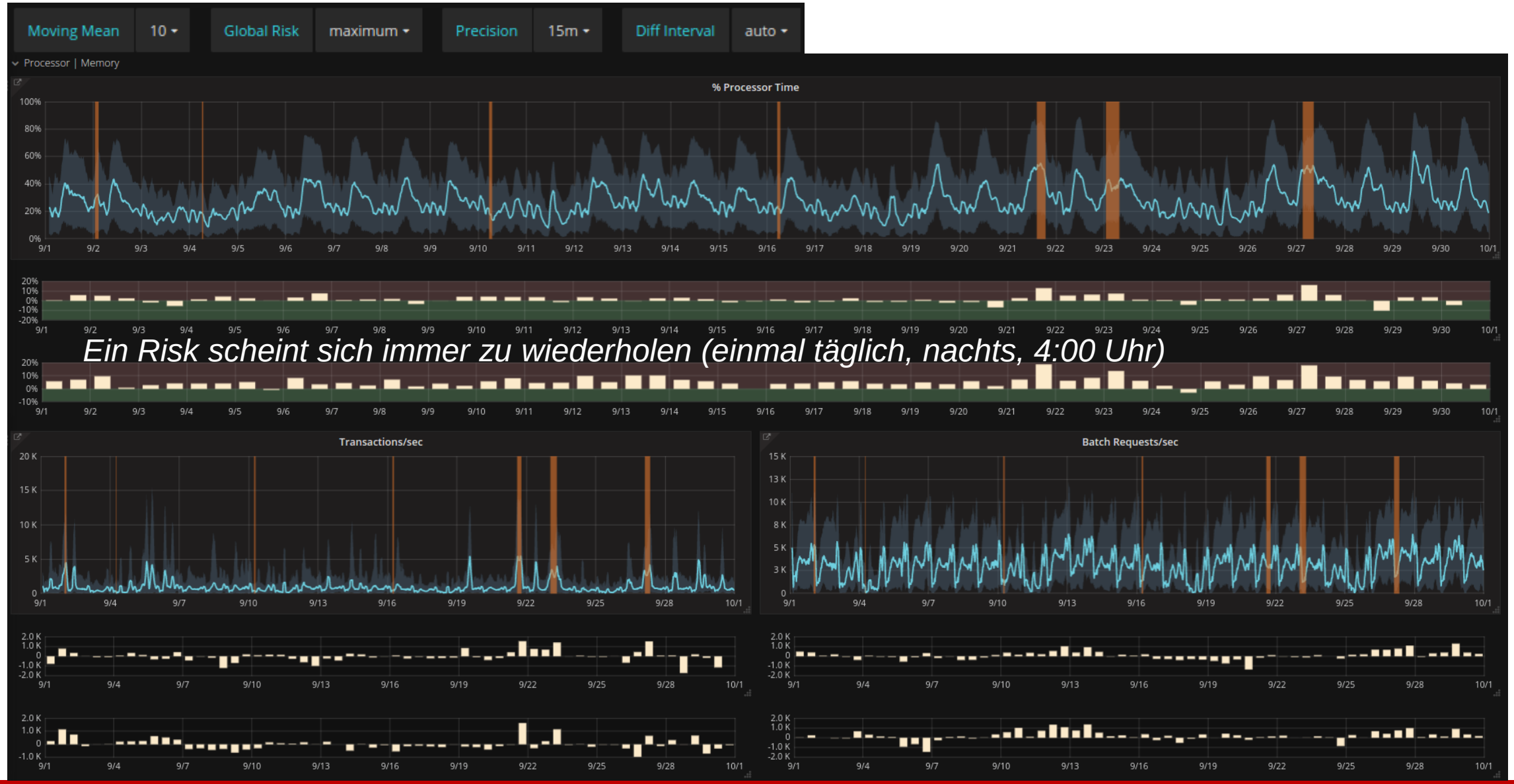
Beispiel I:

Root Cause:
In der für den SQL Server reservierten VM liefen zeitweise auch Transaktionen anderer Prozesse, die eigentlich notwendige Operationen ausbremsten

Die Kombination aus
- *Langzeittrend*
- *Historischen Daten*
- *Risks*
- *Relog & EsxTop Metriken*

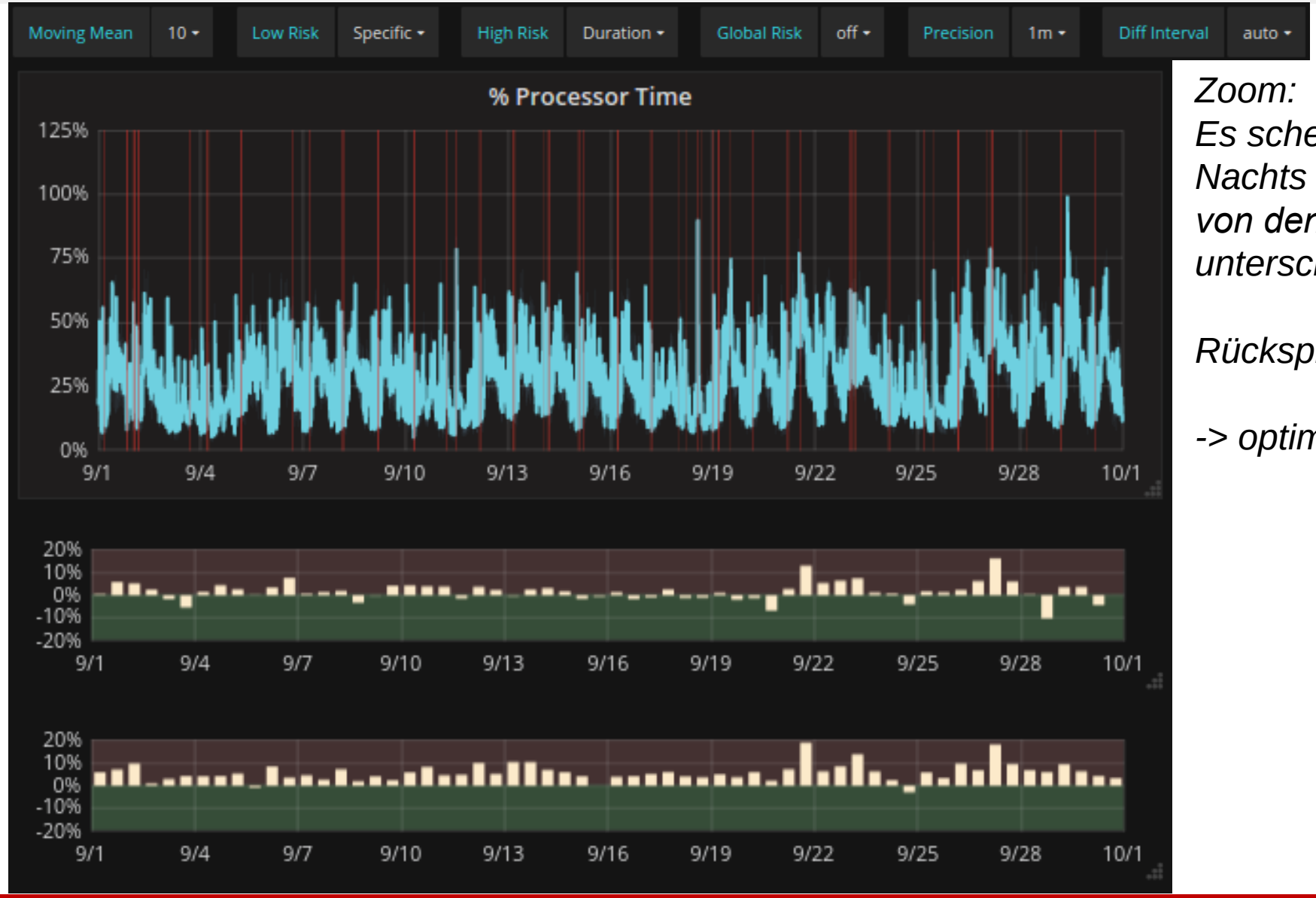
machte die Detektion und Lösung des Problems möglich

Beispiel II: Detektion und Optimierung sich wiederholender Events



Ein Risk scheint sich immer zu wiederholen (einmal täglich, nachts, 4:00 Uhr)

Beispiel II:

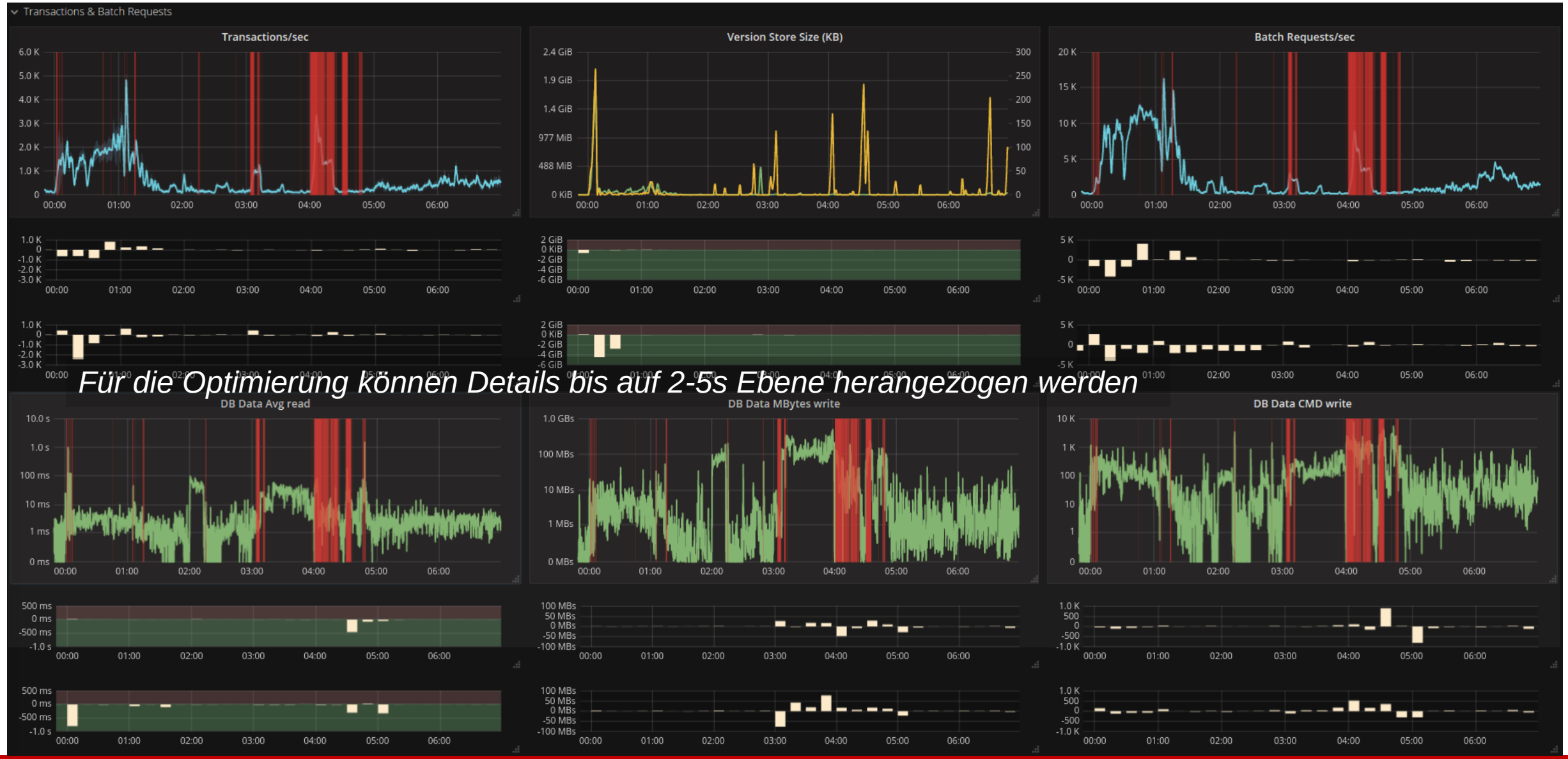


Zoom:
Es scheint jeden Tag zur gleichen Zeit Nachts ein Event stattzufinden, dass sich von den "Standard" Daten merklich unterscheidet.

Rücksprache mit Systemadministrator

-> optimierbarer Cronjob

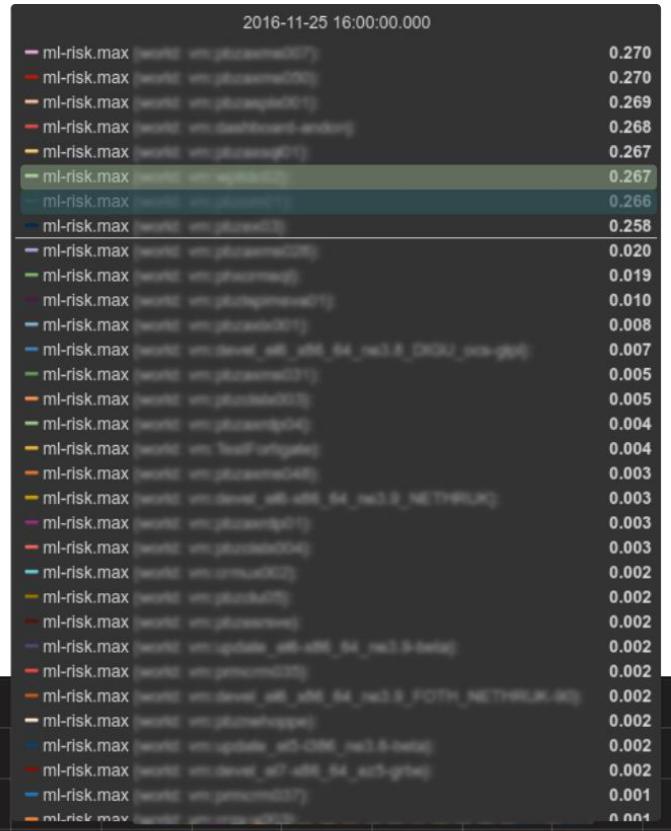
Beispiel II:



Für die Optimierung können Details bis auf 2-5s Ebene herangezogen werden

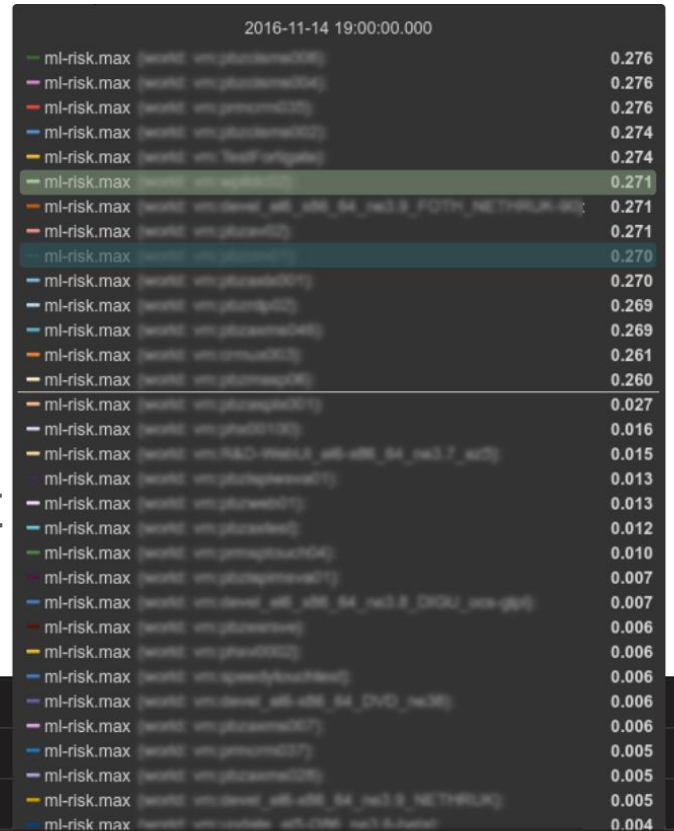
Culprit Detection

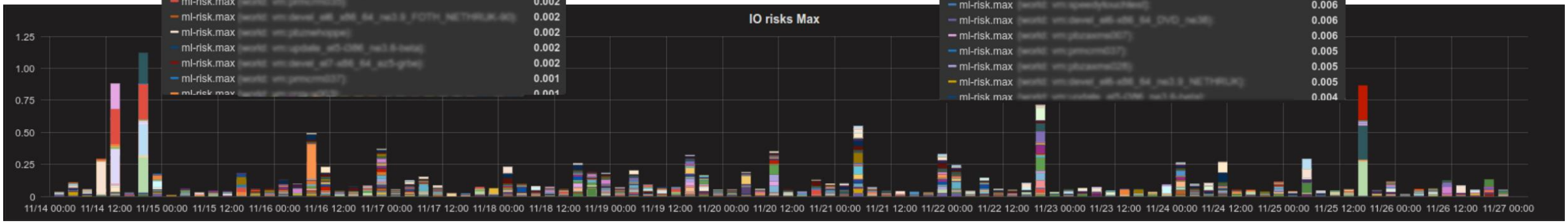
Zu viele virtuelle Maschinen, Kontrolle von Hand unmöglich => check via Risk



Nur wenige VMs haben ein hohes Risk

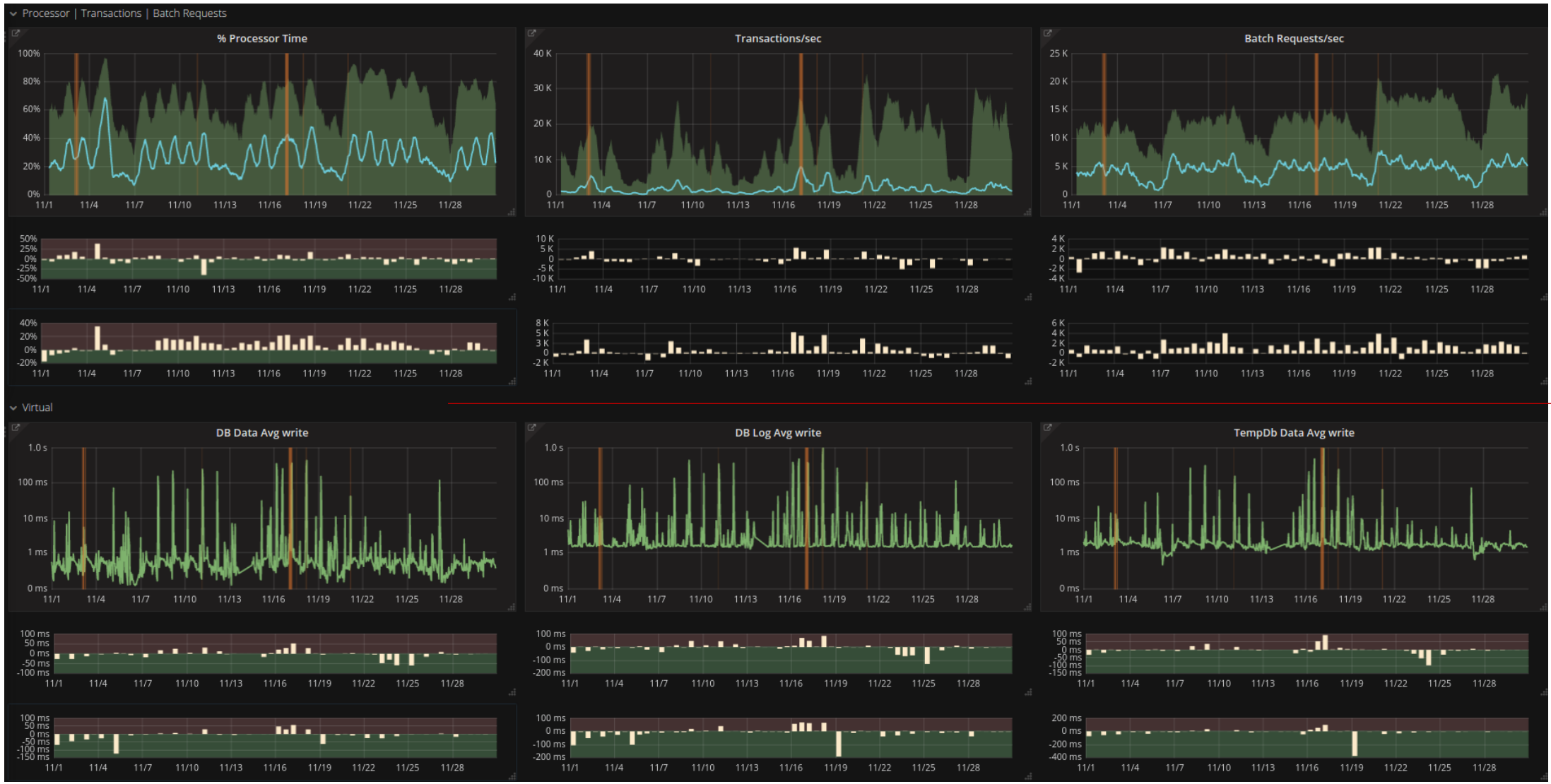
VMs, die mehrfach ein hohes Risk haben sollten zuerst analysiert werden





Zusammenfassung

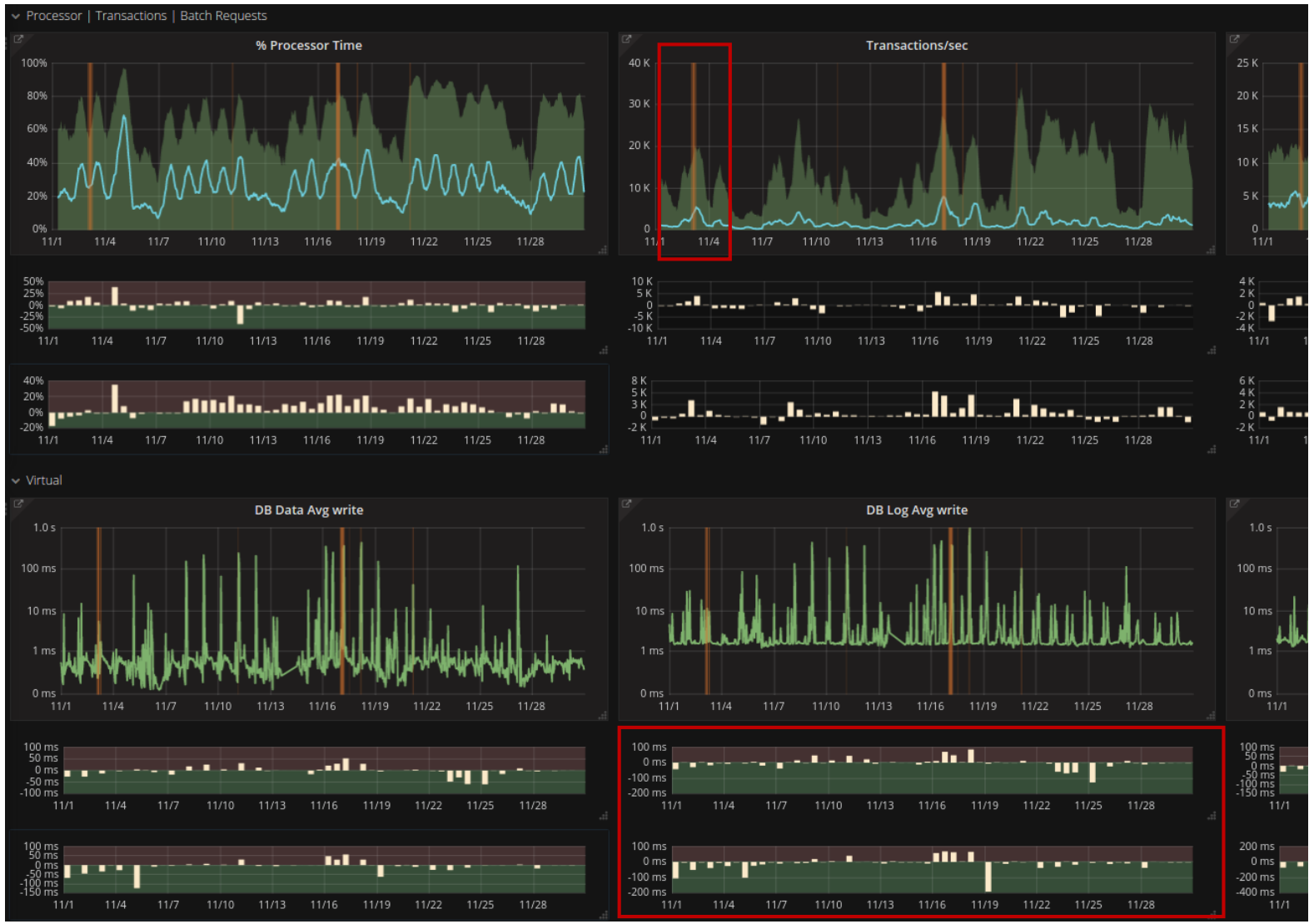
Homogene Ansicht unterschiedlichster Metriken



DATENQUELLE relog

DATENQUELLE esxtop

Risks & Historical Data

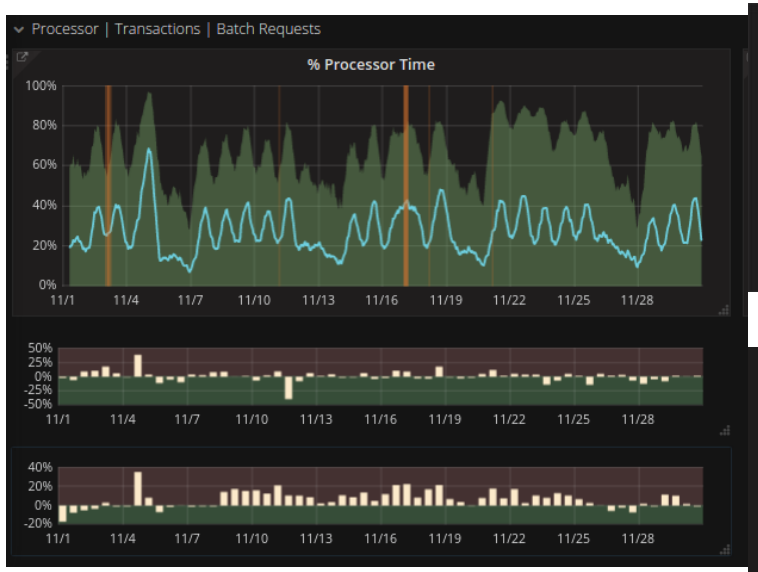


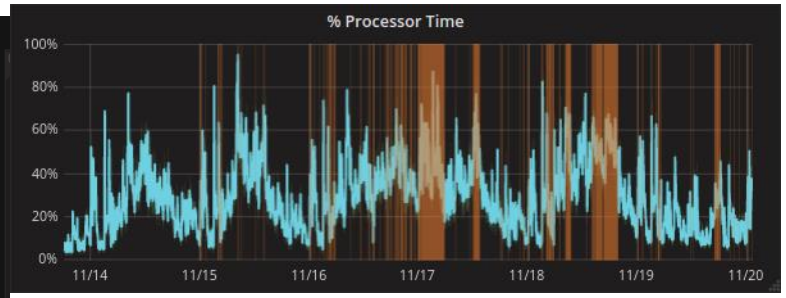
Multivariate RISKs
berechnet aus mehreren Metriken
(alle oder eine Auswahl)

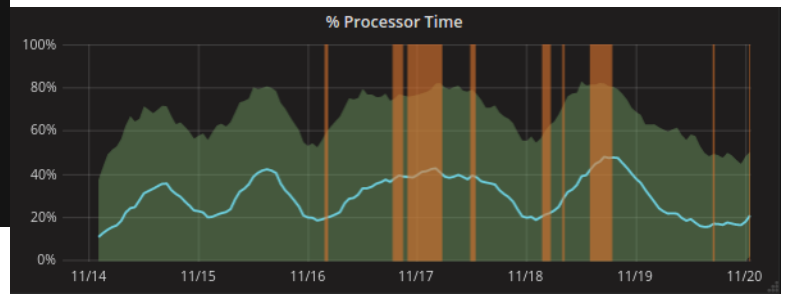
- Schnelleres Troubleshooting
- Mehr Zeit für Root Cause Detection

Historische Daten immer griffbereit
Unterschiede zu den historische Daten
bereits berechnet

Glättung, Bereiche & optimierte Genauigkeit



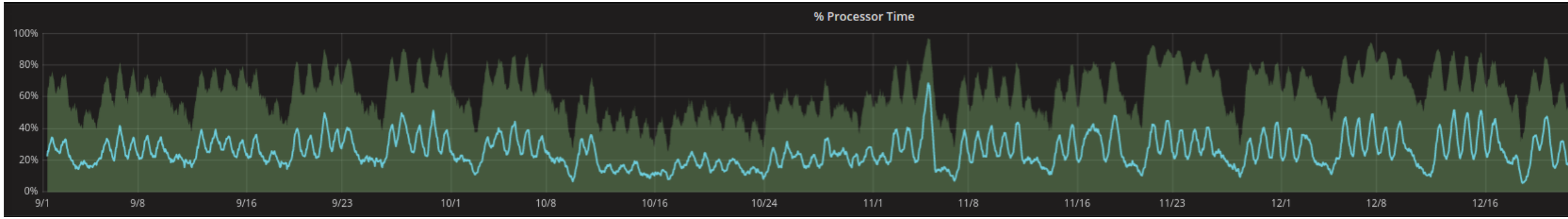




Lade immer nur so viele Daten, wie für die Visualisierung benötigt Multimeasure Zoom

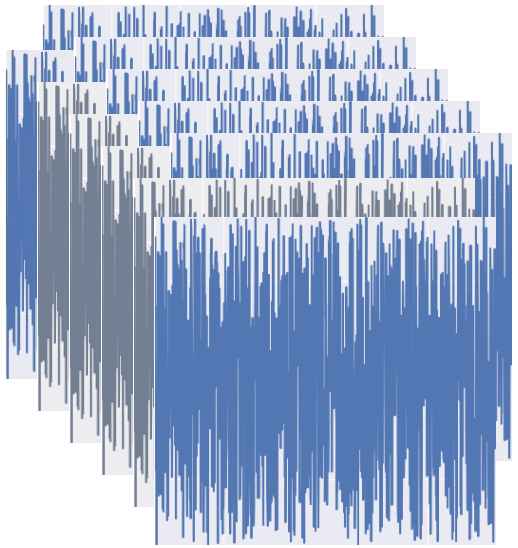
Optimaler Kompromiss: Genauigkeit vs. Speed

Einfache Trenderkennung mit geglätteten Kurven



Risk: Anomaly Detection via Multivariate ML Analysis

Metrics



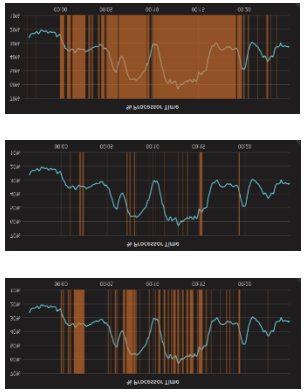
MODELL
STANDARD
VERHALTEN

- Historische Daten der gleichen Metrik
- Historische Daten ähnlicher Metriken
- Historische Daten ähnlicher Maschinen

WAS ERWARTEN WIR?

- Je weiter die Wertekonstellation an einem Zeitpunkt vom Standard entfernt ist, um so höher das Risk
  RISK: anders und selten

WIE WEIT SIND WIR VON UNSERER ERWARTUNG ENTFERNT?



RISK
SCORE

Risk: Culprit Detection mit Risk

IOs

HIGH RISK PERIODS

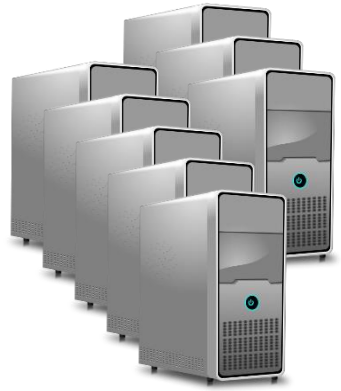
- Historische Daten der gleichen Metrik
- Historische Daten ähnlicher Metriken
- Historische Daten ähnlicher Maschinen
- Historische Daten von Nachbarn

AUTOMATED ANALYSIS

- Welche Maschine(n) ist am wahrscheinlichsten für hohes Risiko verantwortlich
- Proaktive Analysis um Überlastung vorzubeugen

Proaktive Suche von potentiell in der Zukunft gefährlich werdenden Settings

Schnelleres Troubleshooting



- Maschinen mit hohem Risk sollten zuerst untersucht werden, möglicherweise erübrigt sich die Kontrolle der anderen

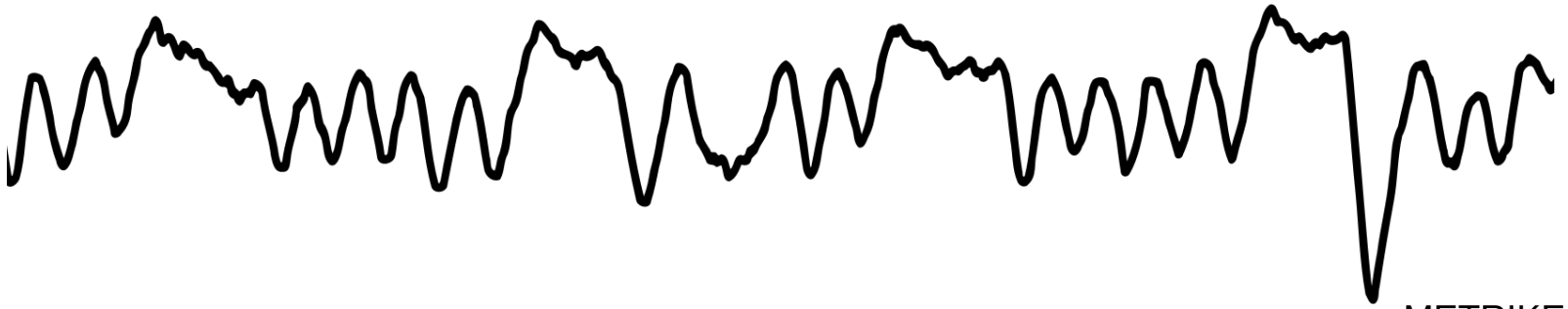
Fazit

Vorteile von Next Level Performance Monitoring

- Risk ist eine nützliche Erweiterung existierender Standard Verfahren
- Weniger, relevantere Alarme
- Situationen die nicht bei Standard Methoden gefunden werden sind detektierbar
- Beziehungen von Metriken untereinander und AD spielen eine Schlüsselrolle für erfolgreiches Next Level Performance Monitoring

- Einsatz menschlicher Experten ist optimiert
    - Kann sich aufs das WAS konzentrieren (und dessen Lösung)
    - Weniger Zeitaufwand, um nach dem WO zu suchen

    => Schnelleres Troubleshooting
    - Proaktives Performance Monitoring (Problem Prävention)



METRIKEN

Abhängigkeiten

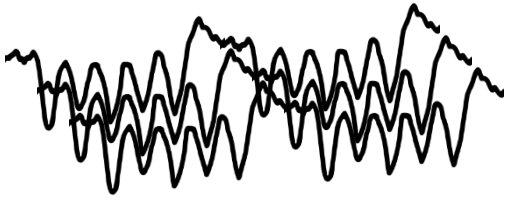
Monitoring & Alarme

Spezifische Daten

Erwartung

- Zeitreihe
- Alarm Schwellwerte

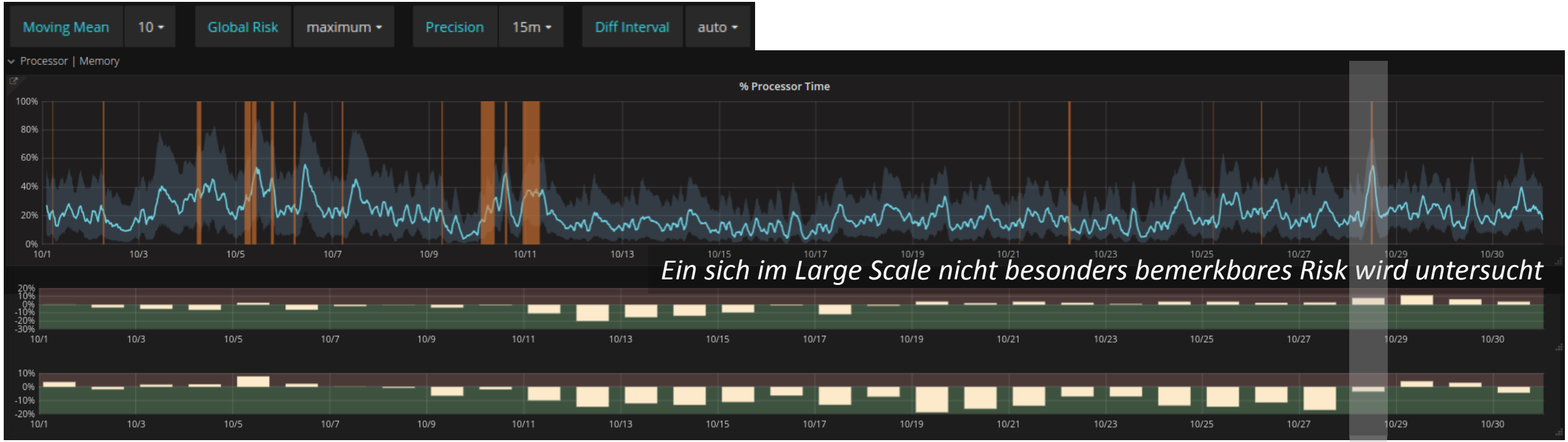
Erfahrung,
ähnliche Daten



Vielen Dank für Ihre Aufmerksamkeit

Supplementary

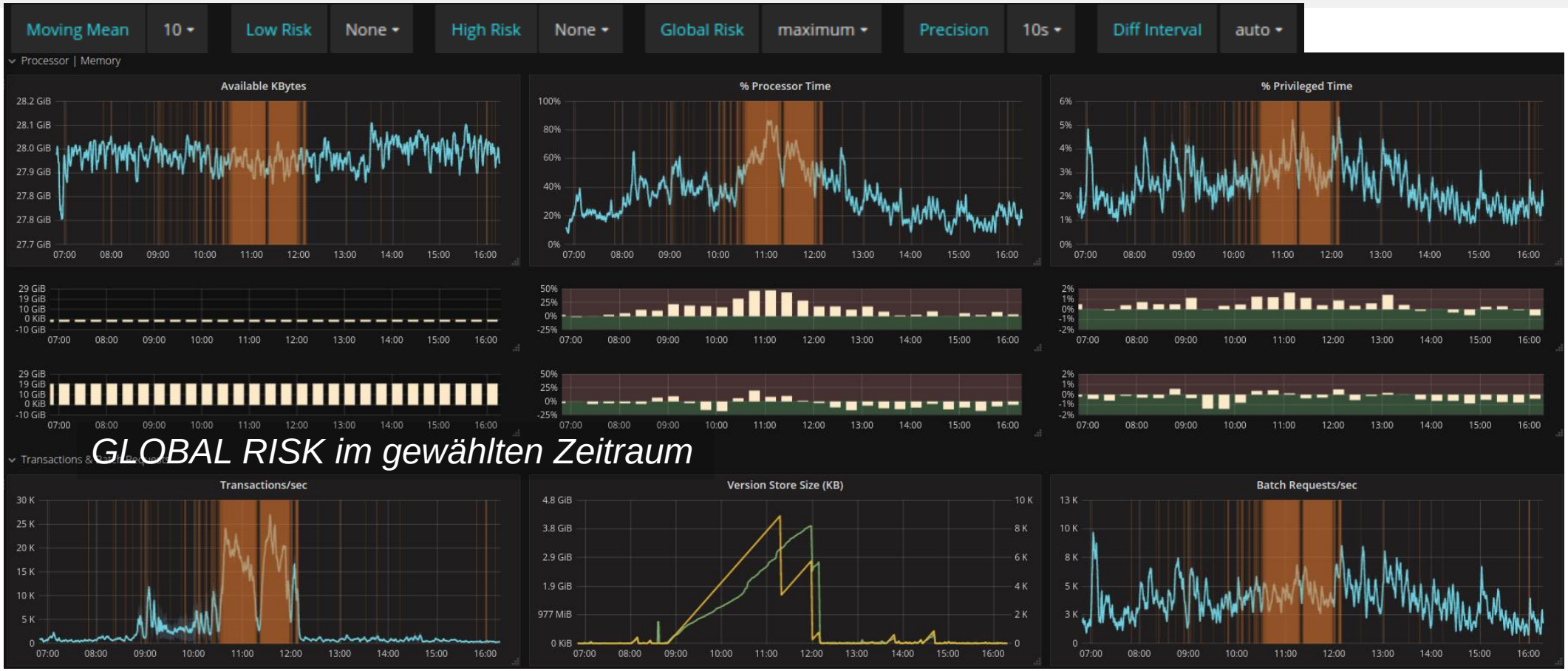
Beispiel III: Minor Event



Ein sich im Large Scale nicht besonders bemerkbares Risk wird untersucht

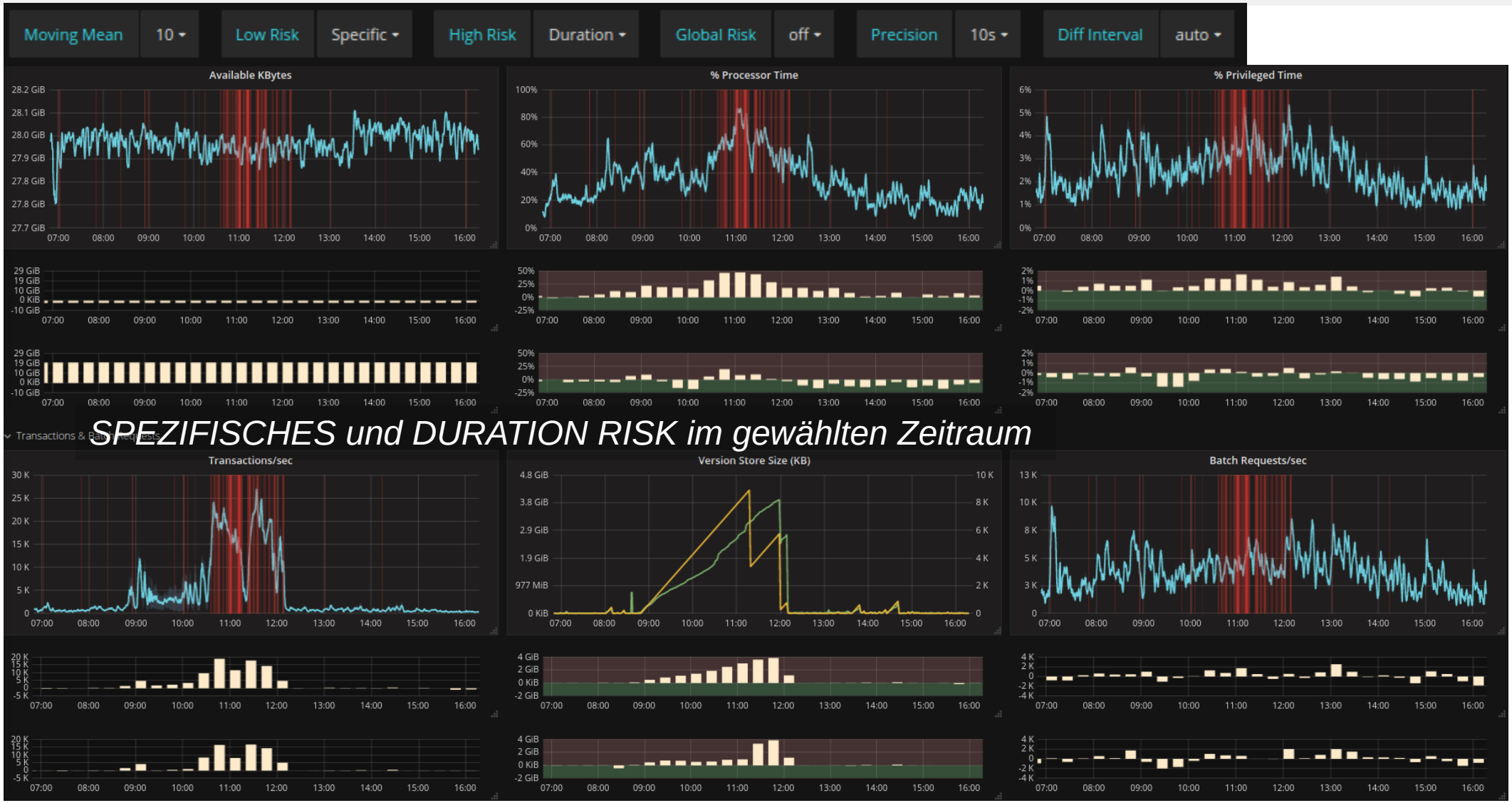
Proaktive Analyse kleinerer Events mit Risk…

Beispiel III: Minor Event



GLOBAL RISK im gewählten Zeitraum

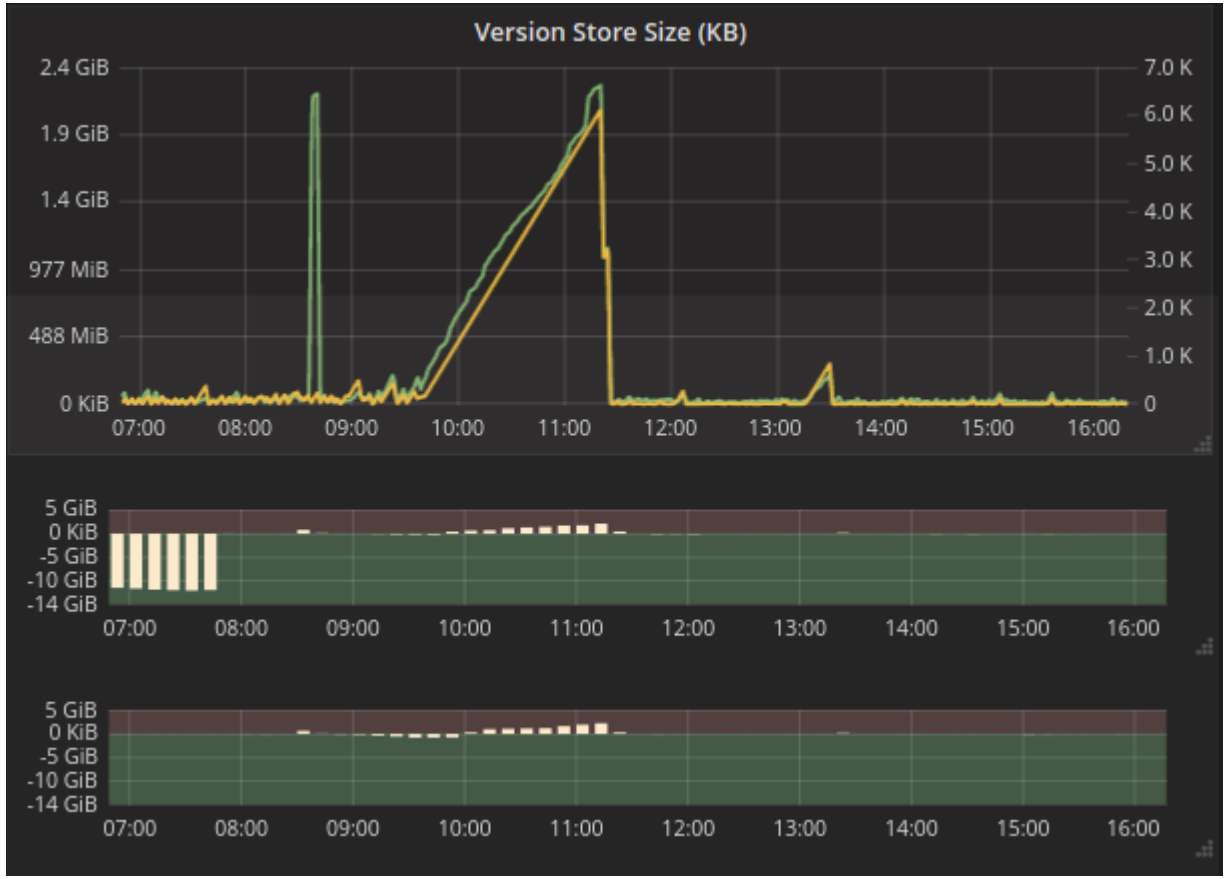
Beispiel III: Minor Event



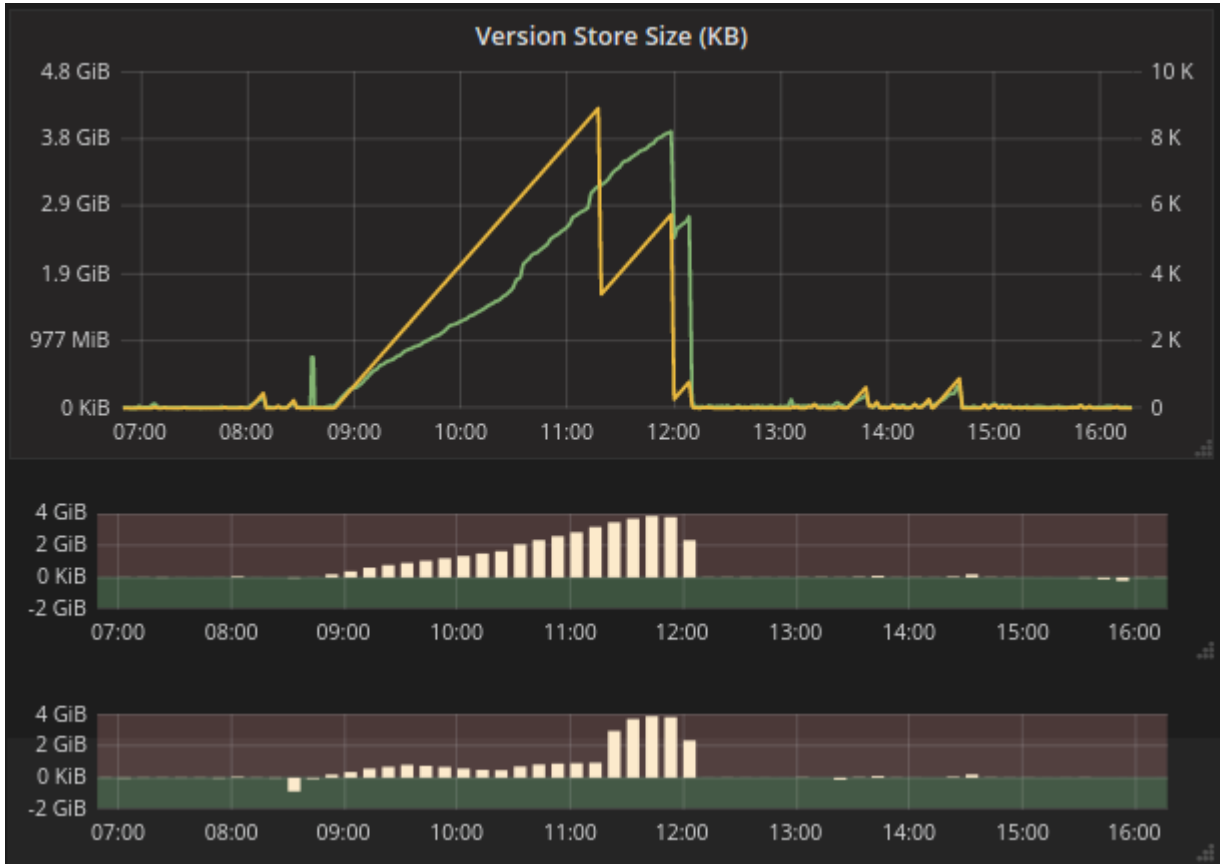
SPEZIFISCHES und DURATION RISK im gewählten Zeitraum

Beispiel III: Minor Event

Vormonat

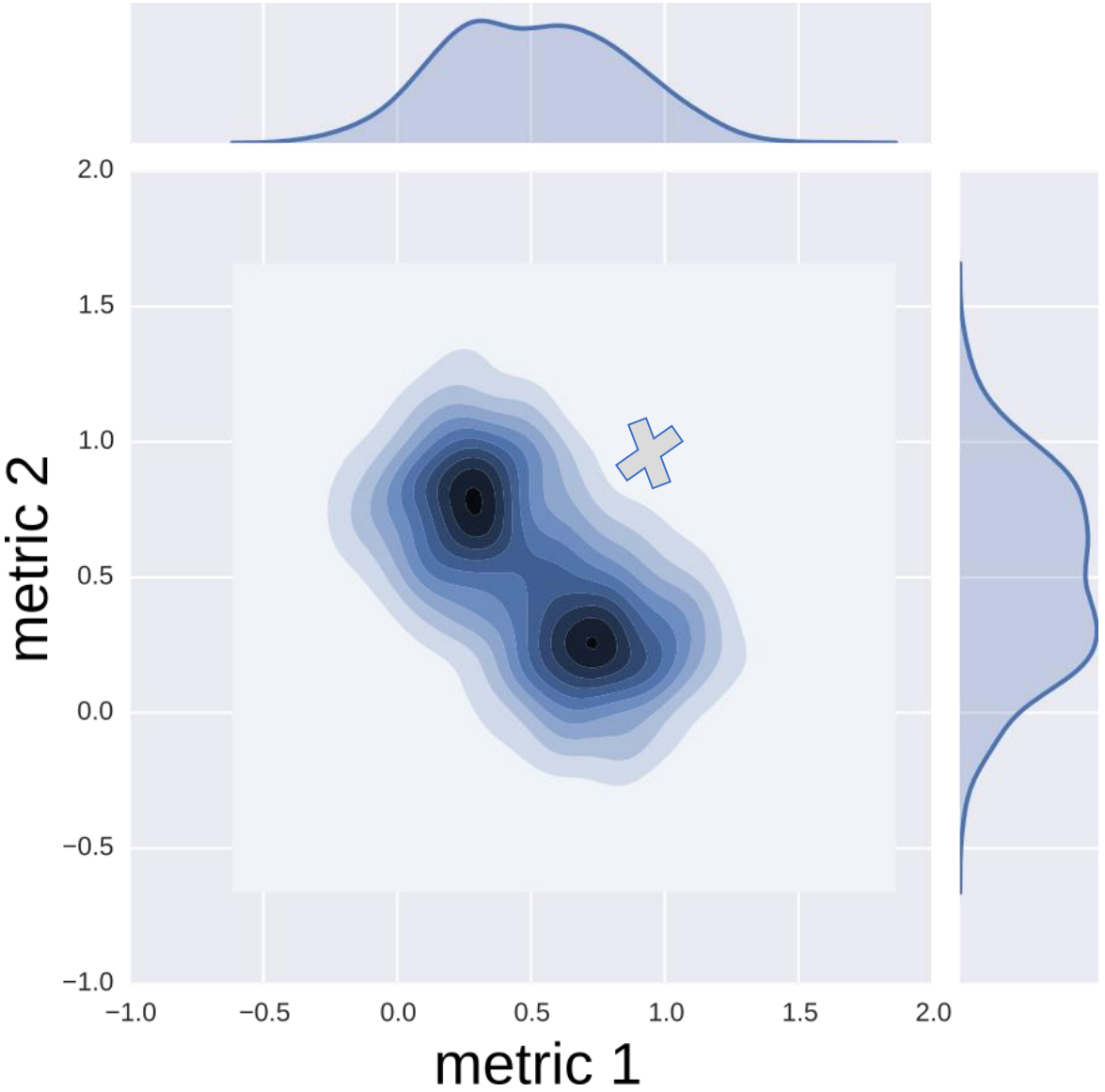


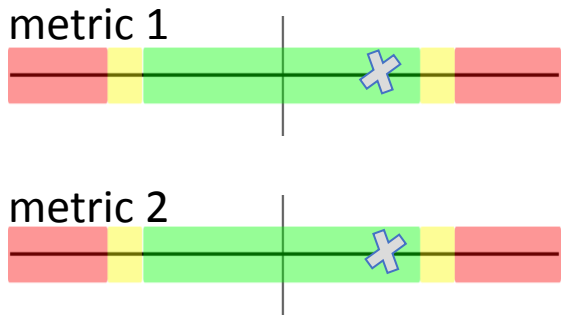
Monat der Analyse



Ein Cronjob war so konfiguriert, dass immer am letzten Freitag des Monats (Mittag 12:00) ein Speichervorgang stattgefunden hat durch zunehmende Daten hätten dadurch negative Auswirkungen entstehen können.
Lösung: Ein anderes Zeitmanagement und Optimierung des Prozesses

Multivariate KPI Analysis





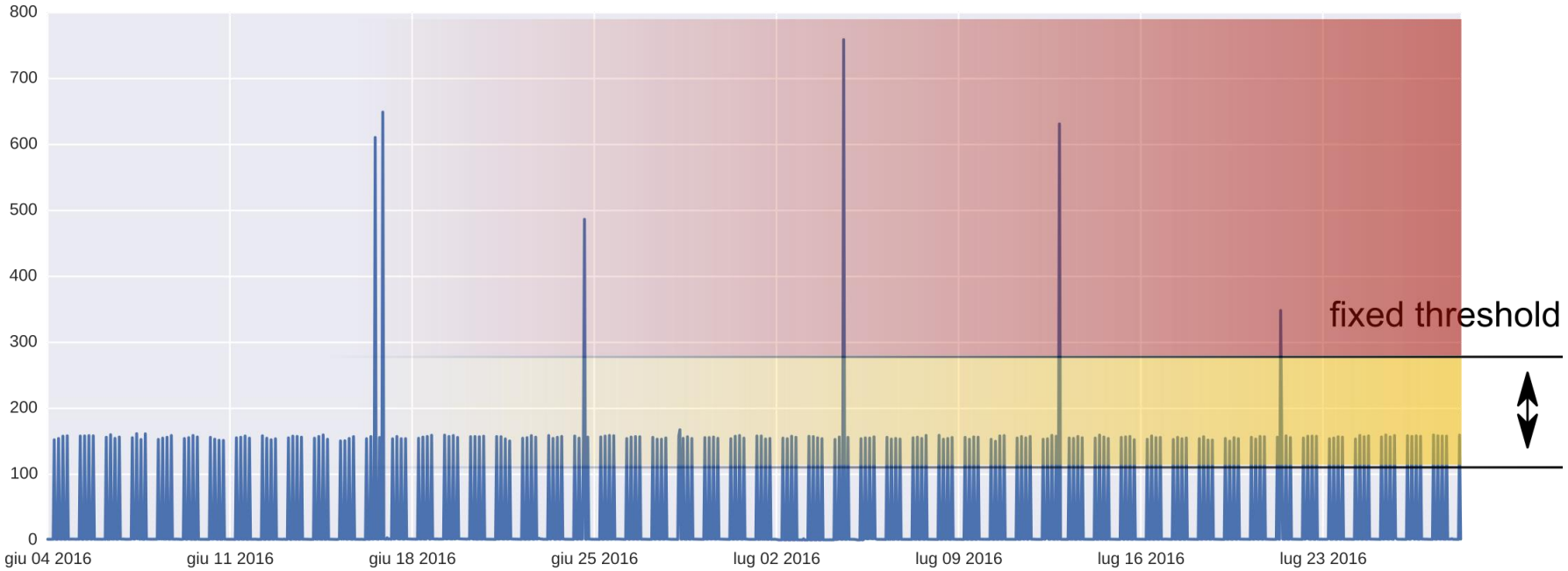
multidimensional data
multivariate methods

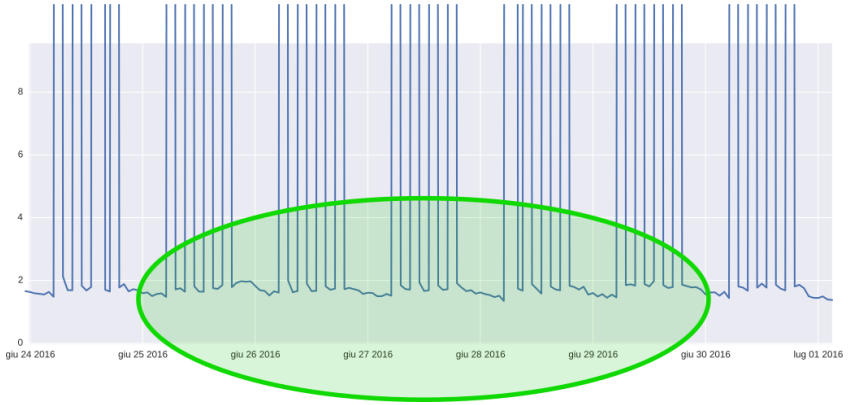
ignoring multidimensionality
=> more false alarms misses

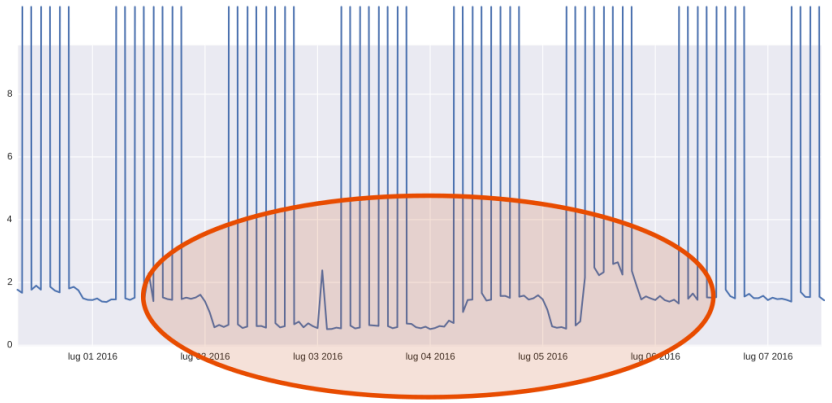
- is close to the max and mean of metric 1
- is close to the max and mean of metric 2
- is not a very probable request in a multidimensional view

➡ a 1D view is not the perfect option for certain types of networks or applications

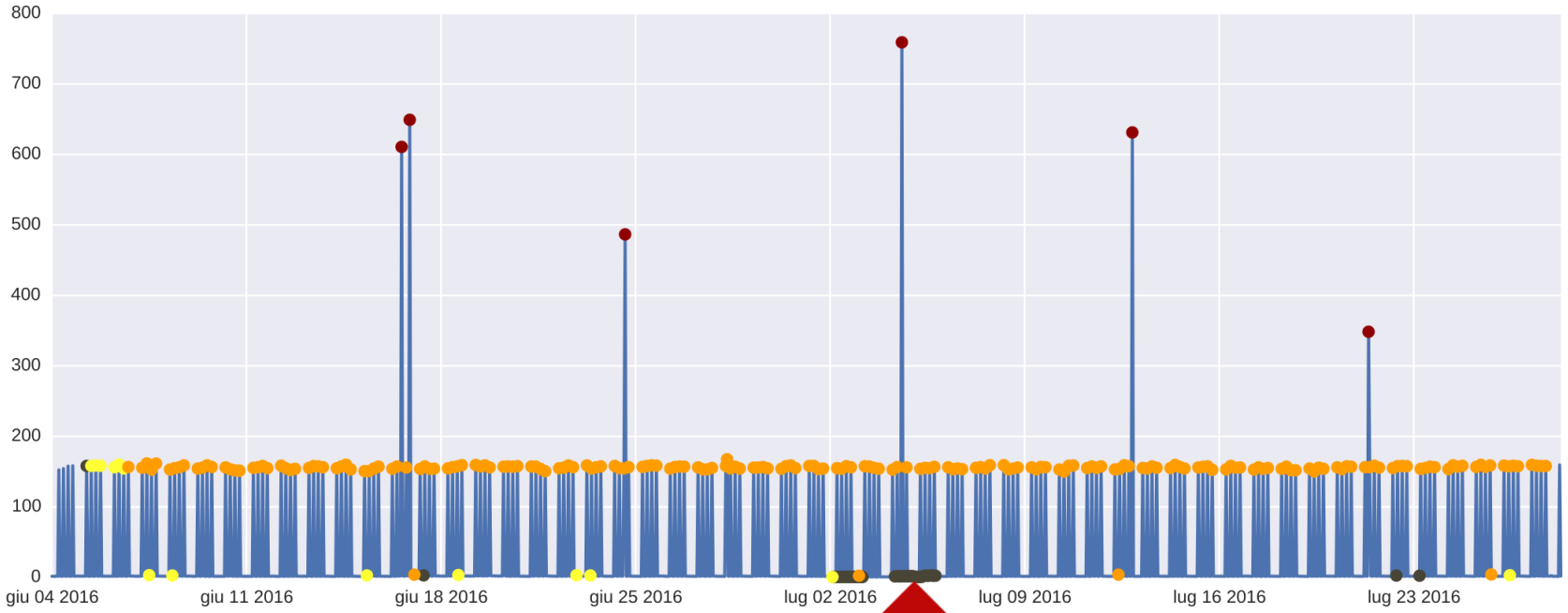
Anomaly Detection

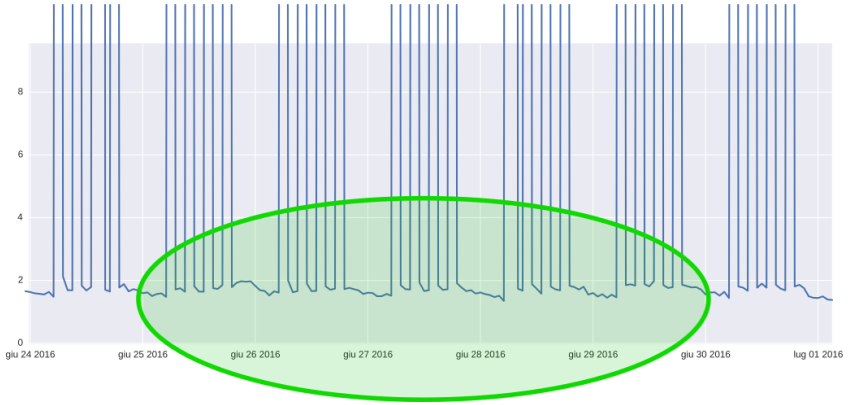


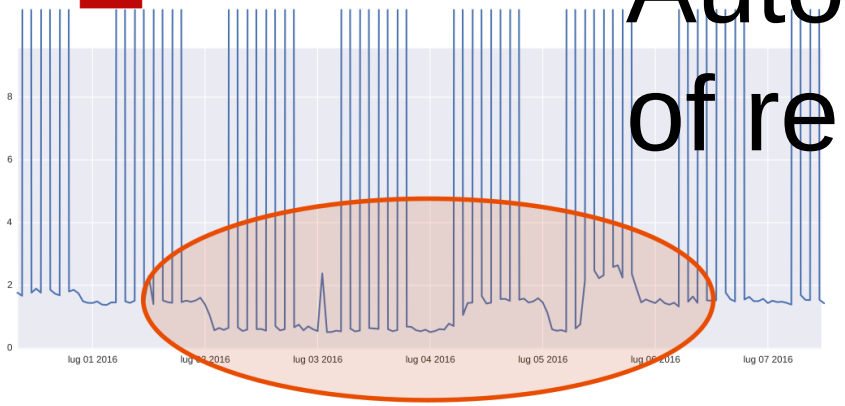




Anomaly Detection

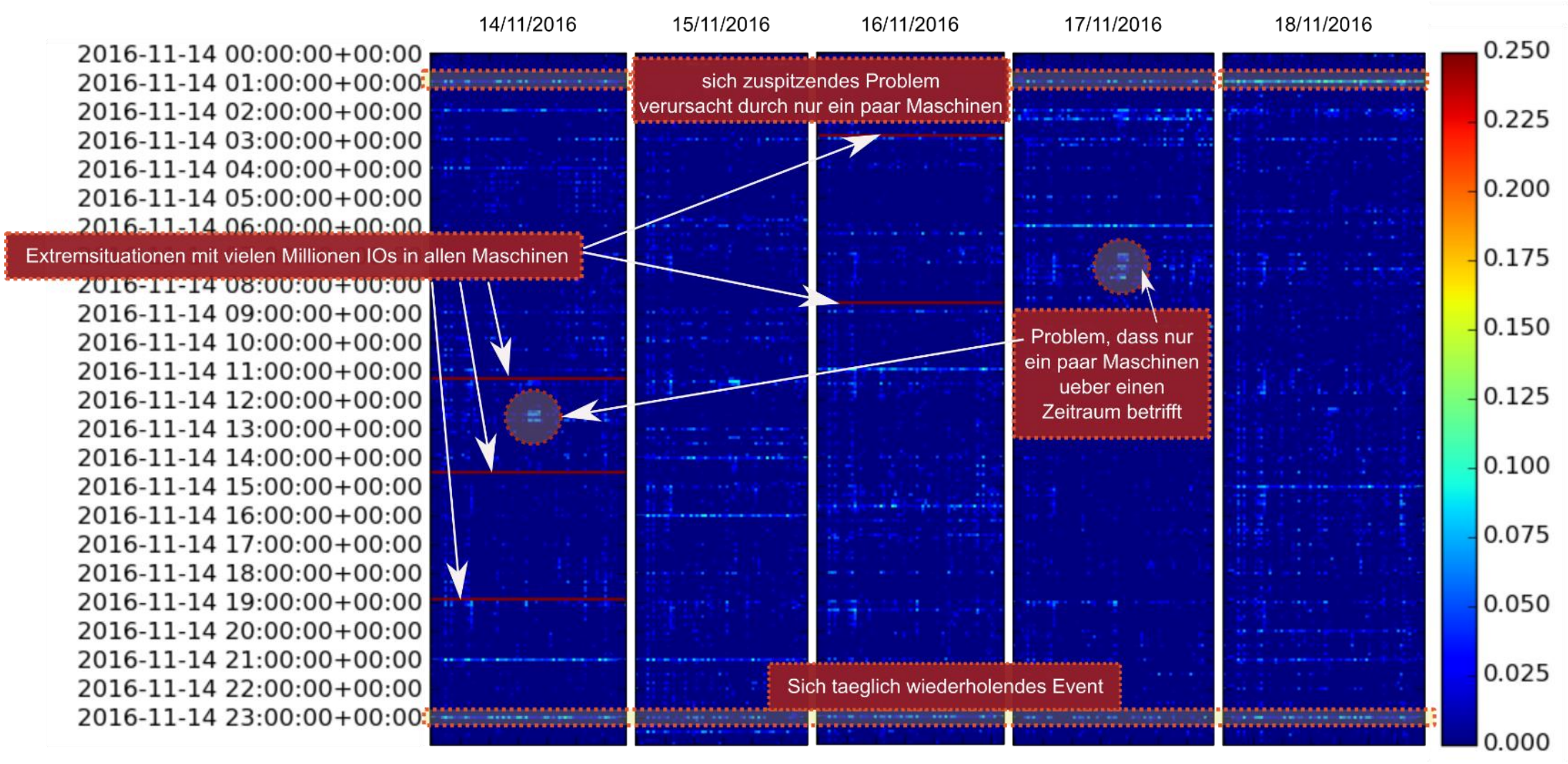




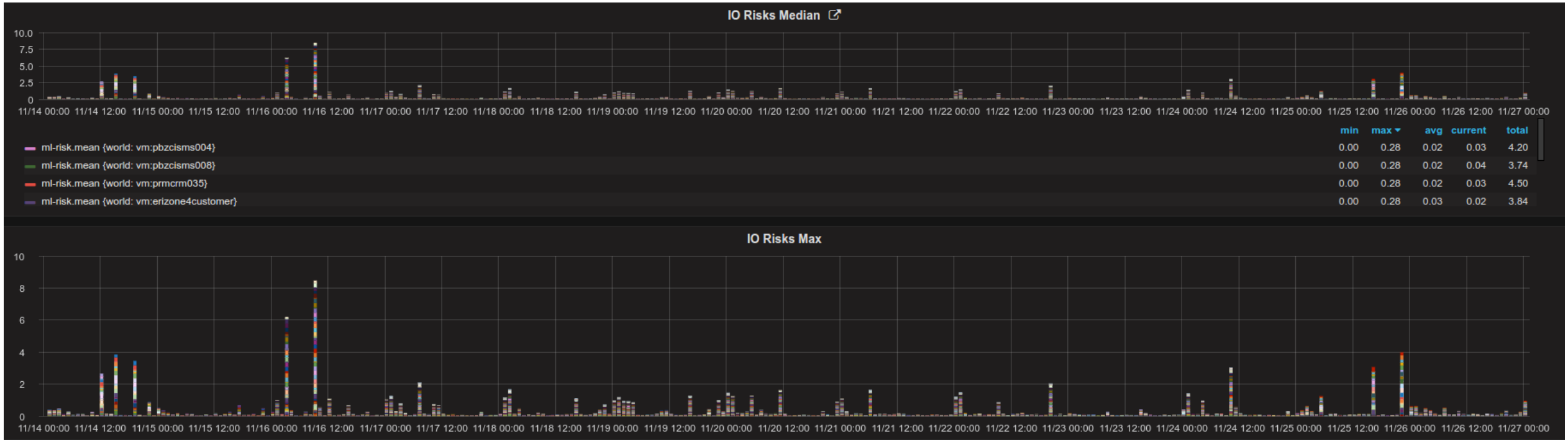


Automatic detection of relevant changes

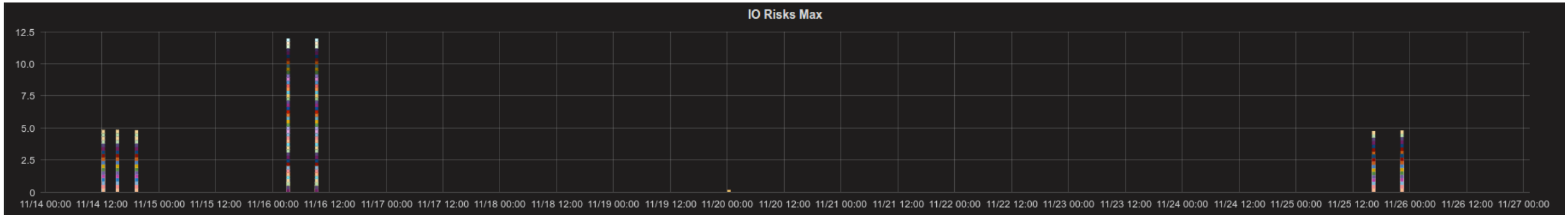
Culprit Detection: Which VM should I blame



Culprit Detection: Which VM should I blame



Threshold on risks not on IO counters



Culprit Detection: Which VM should I blame