



Datawarehousing for Cloud Computing Metrics

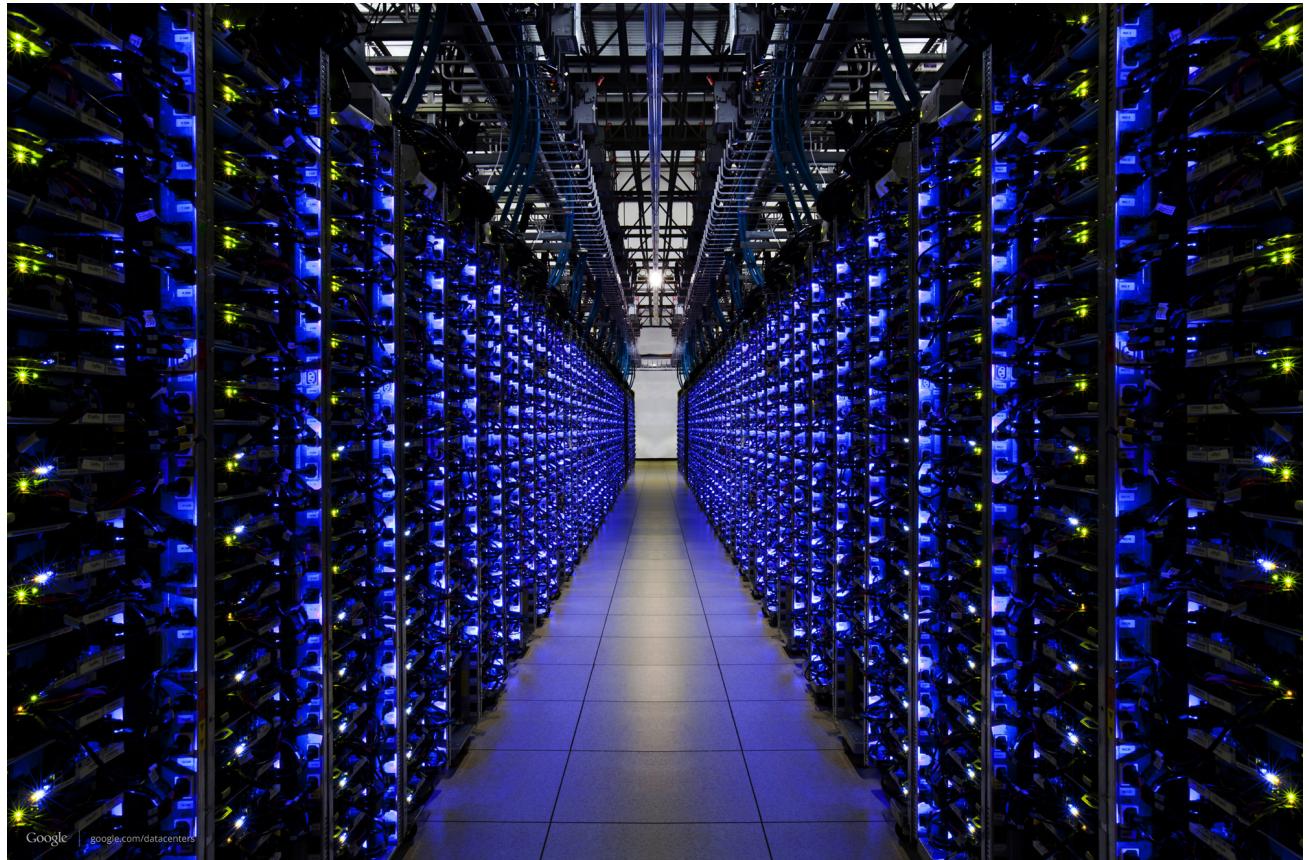
Inhalt

- Einleitung
- Vorstellung - Aktuelle Situation
- Management-Methoden
- Vorbereitung
- Realisierung
- Risiken und unerwartete Probleme
- Abschluss

Einleitung



Einleitung

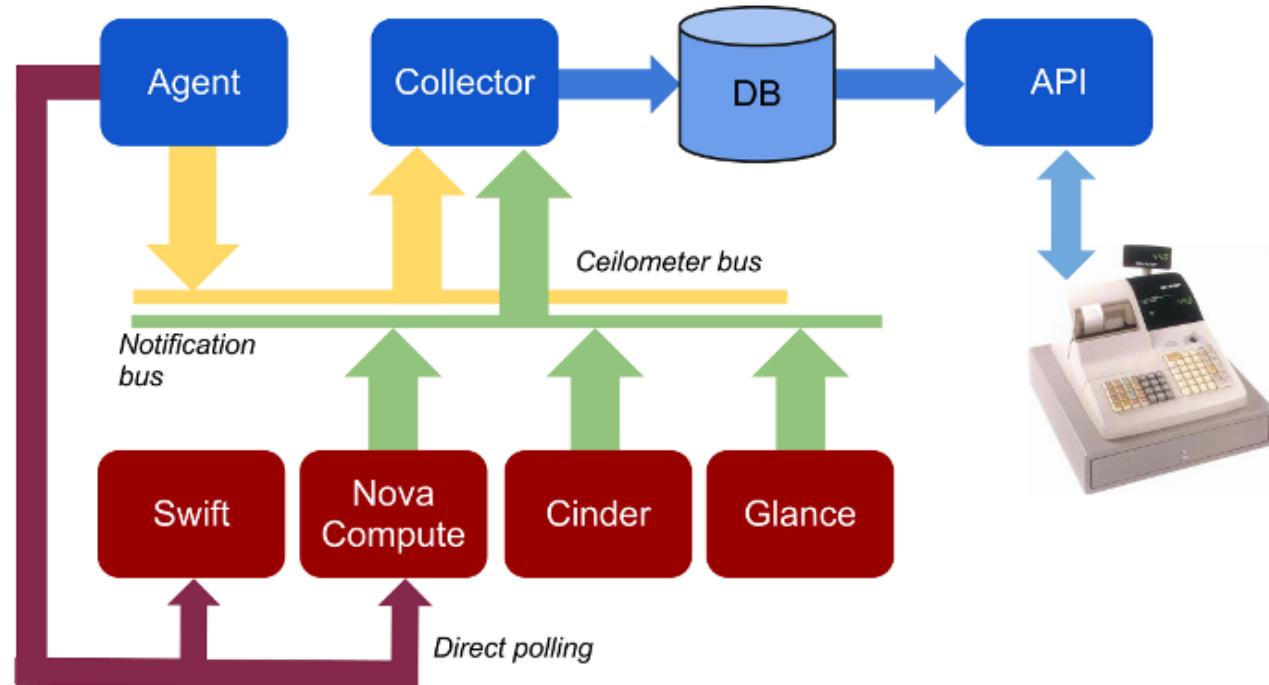


© Google Inc.

Vorstellung



Marktlage



- Telemetry ist ein Teilprojekt von OpenStack
 - Datenbank MongoDB getestet durch Kyle Kingsbury
 - Hohe Hardwareanforderung
 - Benötigt OpenStack Cloud Umgebung

Management

Scrum

- Arbeit in Sprints
- Review & Retrospective
- Beliebt unter Entwicklern
- Sehr starr
- Nicht geeignet für asynchrones Arbeiten

Management

Scrum

Agil

- Flexible Kombination aus Scrum und Kanban
- Nutzung von einzelnen Elementen
 - Sprints
 - R&R
 - Decommissioning

Management

Scrum

Agil

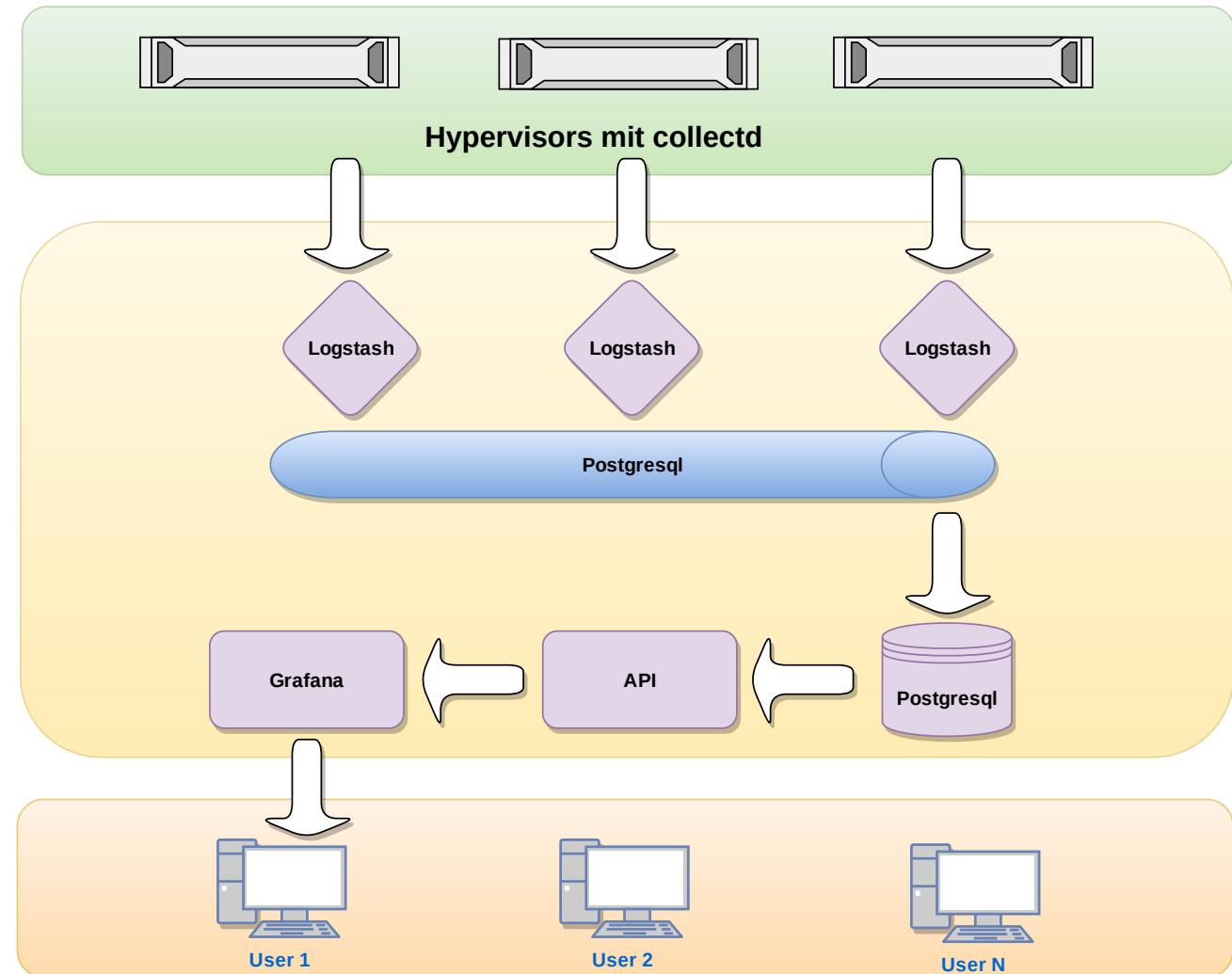
Werkzeuge

- Code Management
 - Git
 - GitHub
 - Continuous Integration: Travis
 - Continuous Delivery: Jenkins
- Projekt Management
 - Jira
 - Confluence
 - Telegram
- Tests
 - Acceptance Tests: Beaker
 - Unit Tests: RSpec
 - Integration Tests: Spring

Vorbereitung

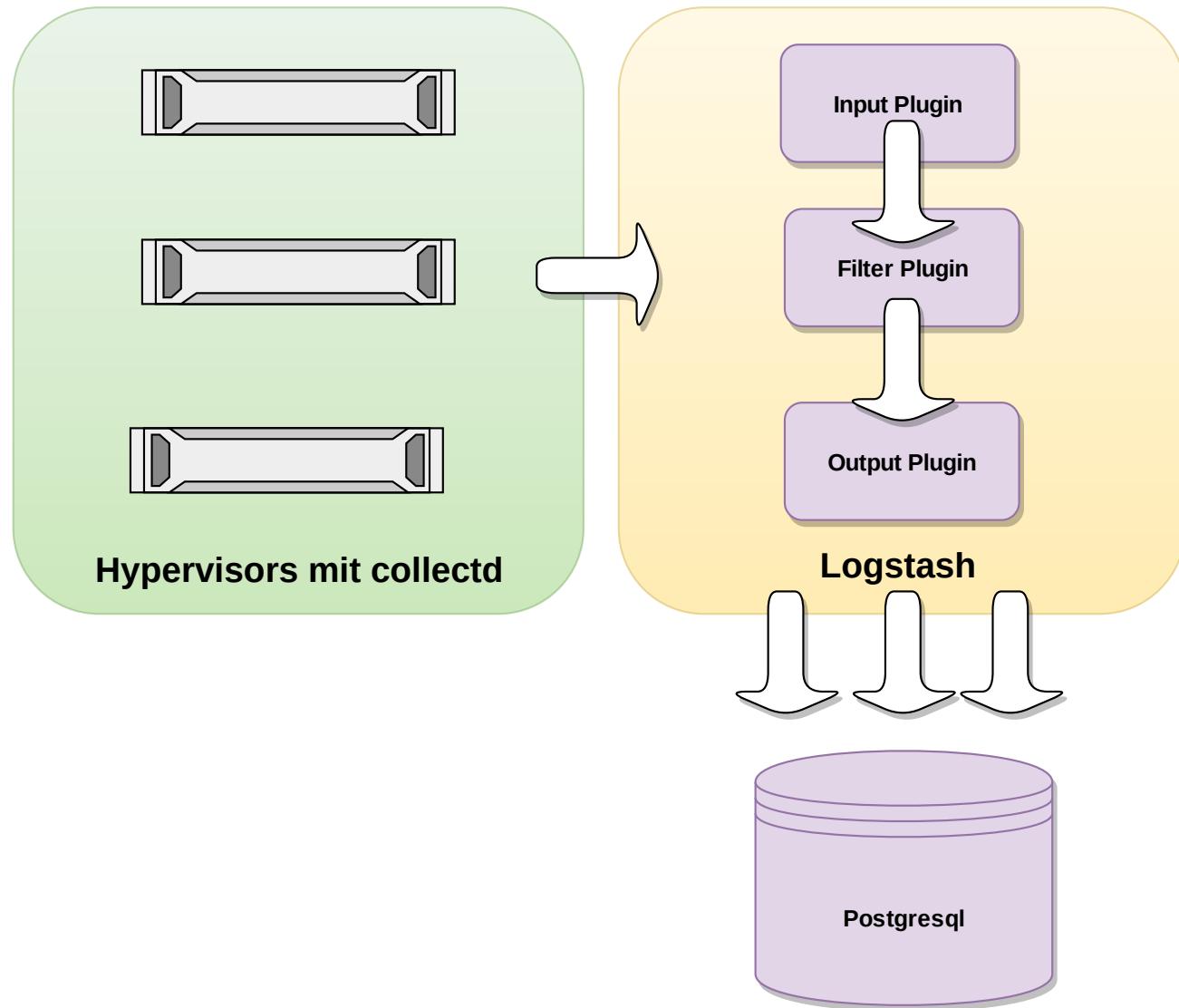
- Absprache mit Auftraggeber bezüglich des Lösungsbildes
- Auswahl von Software zur Realisierung
- Aneignen von Git und Methoden
- Frühzeitig Meilenstein definieren
 - Ersten Prototyp abliefern

Realisierung



Realisierung

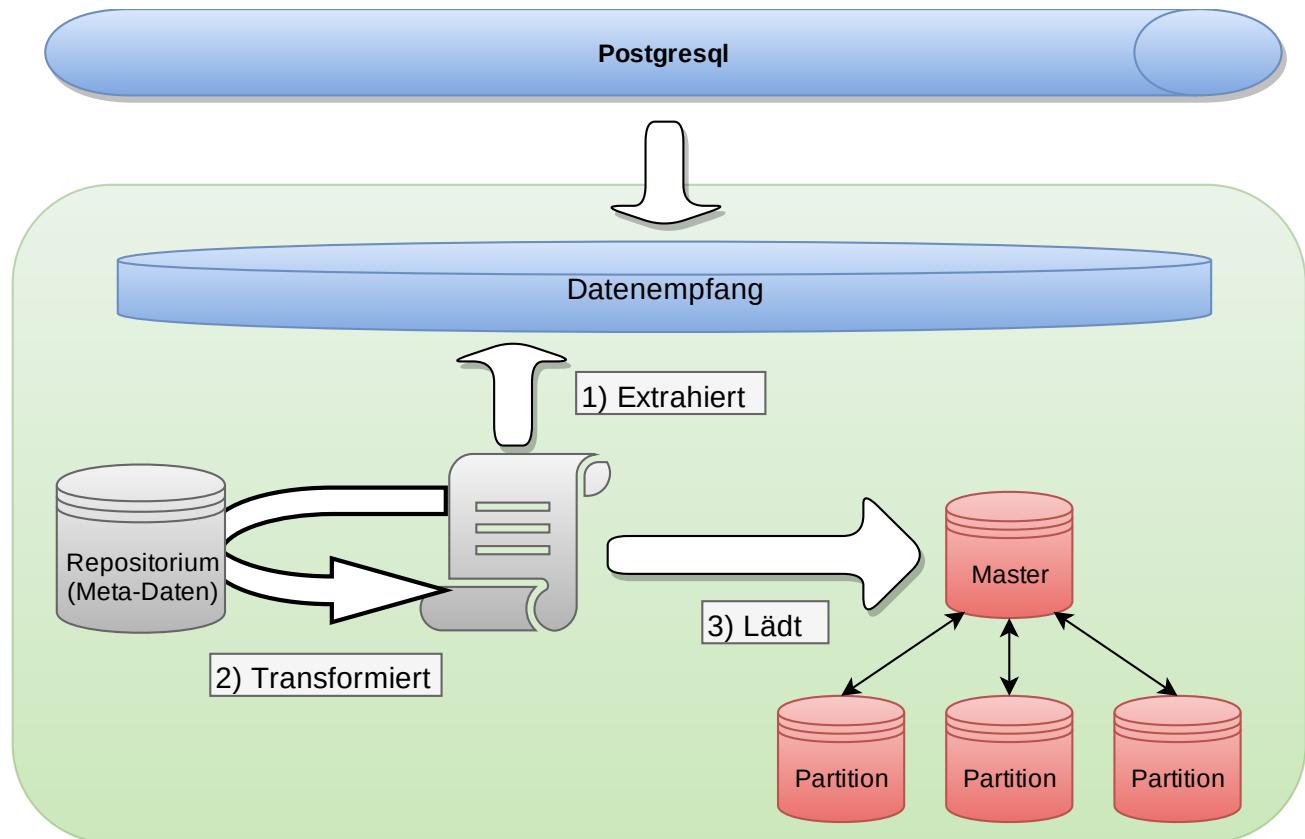
Datenerfassung



Realisierung

Datenerfassung

ETL

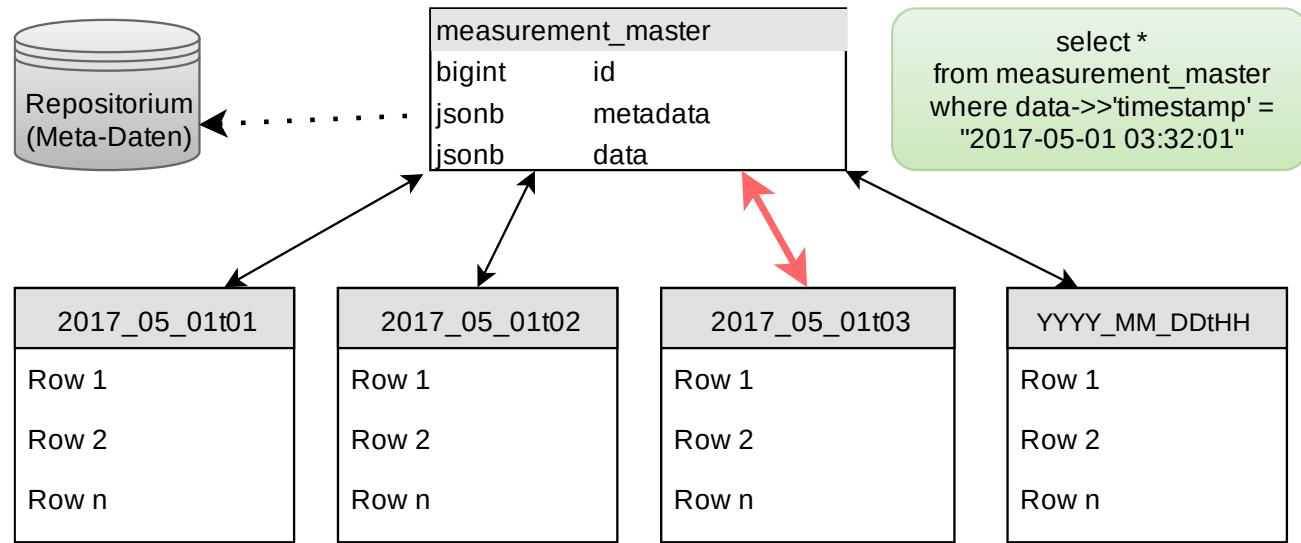


Realisierung

Datenerfassung

ETL

Partitioning



```
{  
  "etlid": "367",  
  "insert_at": "2017-0501T03:32:28.575413",  
  "aggregation_type": "seconds"  
}
```

```
{  
  "host": "vm22062017",  
  "value": "10",  
  "plugin": "ram",  
  "timestamp": "2017-05-01 03:32:01",  
  "collectd_type": "percent",  
  "type_instance": "used"  
}
```

Realisierung

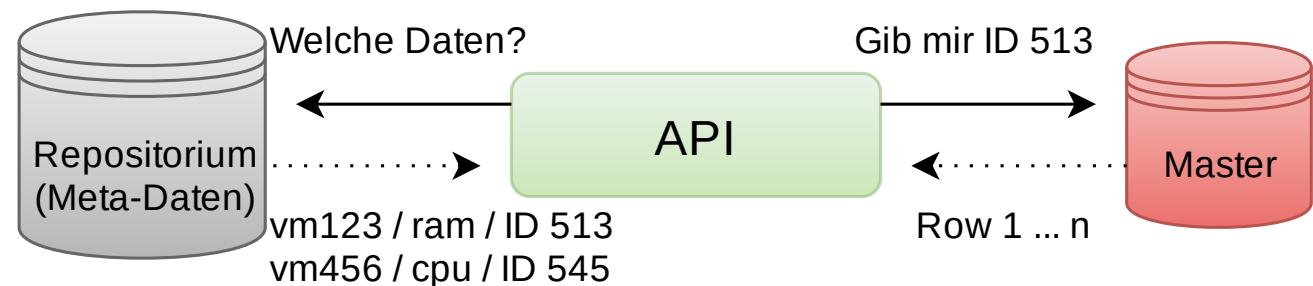
Datenerfassung

ETL

Partitioning

API

- Bereistellung von Funktionen
 - Datenverarbeitung muss nicht in SQL erfolgen
 - Vorgabe der Daten durch die API
 - Keine Notwendigkeit für direkte SQL Queries
- Zusätzliche Abstraktionsschicht
 - Zusätzlicher Cache
 - Datenbankaustausch ist transparent für Grafana
- Kleinere Angriffsfläche auf das Interface
 - HTTP API vs SQL
 - Einfacher zu testen



Realisierung

Datenerfassung

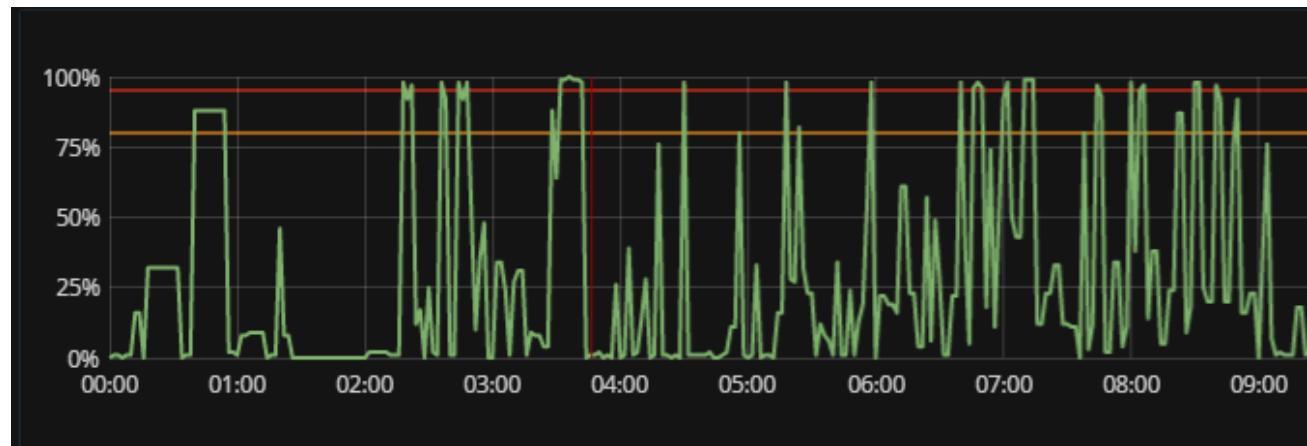
ETL

Partitioning

API



Visualisierung



Realisierung

- SSD Userstory
- CPU Userstory
- Memory Userstory
- Zeitdefinierte Analyse Userstory
- Webinterface Userstory

ETL

Partitioning

API

Visualisierung

Userstorys

Realisierung

Datenerfassung

ETL

Partitioning

API

Visualisierung

Userstorys

Wireframe

Risiken

- Risk Register definieren
- Präventivmaßnahmen
- Post Mortem Maßnahmen

Risiken



Ausblick

- Alerting
 - Benachrichtigung bei Schwellenwertüberschreitung
- Quality of Service
 - Bevorzugen bestimmter Datentypen zur Live-Visualisierung
- Cache Invalidation
 - Sicherstellung der kontinuierlichen Datenverarbeitung
- Trendgenerierung
 - Extrapolation vorhandener Daten

Fazit

Solide und modulare Architektur

Flexibler und einfacher als Telemetry

Persönliche Fortbildung