

Schriftliche Zusammenfassung des Vortrags im Oberseminar

„Datenbanksysteme – Aktuelle Trends“

im Sommersemester 2019

bei Herrn Prof. Dr. Thomas Kudraß

Thema:

Time Series Databases – Zeitreihendatenbanken

Bastian Walthierer

Matrikelnummer: 72860

Inhaltsverzeichnis:

Mathematische Begriffserklärung	3
Zeitreihen	3
Zeitreihenanalyse	3
Zeitreihendatenbanken.....	3
Eigenschaften:	4
Verschiedene Datenmodelle	5
Umgang mit Datenmengen	5
Arten von TSDBs	6
Verbreitung von TSDBs.....	6
Anwendung von TSDB.....	6
Demo InfluxDB.....	7
Fazit	7
Quellenangaben:.....	8

Mathematische Begriffserklärung

Zeitreihen

Erklärung des Begriffes der Zeitreihen

Zeitreihenanalyse ¹

Definition des Begriffes Prognose und Herstellung des Zusammenhangs mit Zeitreihenanalyse.

Bestandteile einer Zeitreihenanalyse.

Ziel der Zeitreihenanalyse.

Zeitreihendatenbanken

Entstehung: ²³

Gründe für die Entstehung von Zeitreihendatenbanken, wie z. B.

Bedürfnis große Datenmengen, die in Abhängigkeit der Zeit, generiert werden in Datenbanken abzuspeichern, Sensordaten, Überwachung der Finanzmärkte, Herausforderungen mit bereits bekannten Datenbanksystemen.

¹ vgl. Einführung in die Zeitreihenanalyse, Kreiß, J. P. K., & Neuhaus, G. N. (o.D.). Google Books. Abgerufen 15. Juli, 2019, von <https://books.google.de/?hl=de>

² vgl. Bader, A. B., Kopp, O. K., & Falkentahl, M. F. (2017). Survey and Comparison of Open Source Time Series Databases. Abgerufen 15. Juli, 2019, von ftp://ftp.informatik.uni-stuttgart.de/pub/library/ncstrl.ustuttgart_fi/INPROC-2017-06/INPROC-2017-06.pdf

³ vgl. Fangman, S. F. (2019, 28. Januar). Data Driven Investor. Abgerufen von <https://medium.com/datadriveninvestor/the-time-has-come-for-a-new-type-of-database-47cf8df1667a>

Anforderungen:⁴⁵

Liste von Anforderungen an TSDBs:

- Time Series
- In-memory
- Schneller Zugriff auf Datensätze
- Hohe Kapazitäten

Eigenschaften:⁶⁷⁸

Bestandteile

- Time Stamp
- Time Range
- ID und mehreren Rows
- Row besteht aus Time Stamps, Values, optionalen Namen / Werten und Indexen

Funktionen

⁴ vgl. Time Series Databases explained. (o.D.). Abgerufen von <https://www.influxdata.com/time-series-database/>

⁵ vgl. Bader, A. B., Kopp, O. K., & Falkentahl, M. F. (2017). Survey and Comparison of Open Source Time Series Databases. Abgerufen 15. Juli, 2019, von ftp://ftp.informatik.uni-stuttgart.de/pub/library/ncstrl.ustuttgart_fi/INPROC-2017-06/INPROC-2017-06.pdf

⁶ Key/Value-Datenbanksysteme. (o.D.). Abgerufen von <http://www.datenbanken-verstehen.de/lexikon/key-value-datenbanksysteme/>

⁷ vgl. Bader, A. B., Kopp, O. K., & Falkentahl, M. F. (2017). Survey and Comparison of Open Source Time Series Databases. Abgerufen 15. Juli, 2019, von ftp://ftp.informatik.uni-stuttgart.de/pub/library/ncstrl.ustuttgart_fi/INPROC-2017-06/INPROC-2017-06.pdf

⁸ vgl. Time Series Databases explained. (o.D.). Abgerufen von <https://www.influxdata.com/time-series-database/>

Verschiedene Datenmodelle⁹¹⁰

- Datenmodell RDBMS mit Time Series Data
Speicherung der Daten mittels Sternschema.
- Datenmodell NoSQL DBMS
Speicherung der Daten mittels Objekte, Dokumente oder Listen.
- Datenmodell NoSQL mit JSON Blob
Tabellen werden durch das Zusammenfügen aller Daten einer Row in eine einzige Datenstruktur namens Blob verbessert.
- Datenmodell TSDB mit Tagsets
Alle Datenmengen besitzen eine Spalte Time in der Timestamps mit Datum und Uhrzeit gespeichert werden. Zusätzlich gibt es Spalten mit Fields, die wiederum aus Field Keys und Field Values bestehen. Zudem gibt es auch noch Tags. Hier wird auch wieder unterschieden in Tag Keys und Tag Values.

Umgang mit Datenmengen¹¹

- Skalierbarkeit
ACID Eigenschaften
- Relevante Daten
- Downsampling / Gruppierung

⁹ Time Series Databases: New Ways to Store and Access Data. (2014). Abgerufen von <https://apprize.info/data/series/4.html>

¹⁰ vgl. InfluxDB key concepts. (o.D.). Abgerufen von https://docs.influxdata.com/influxdb/v1.7/concepts/key_concepts/

¹¹ vgl. Fangman, S. F. (2019, 28. Januar). Data Driven Investor. Abgerufen von <https://medium.com/datadriveninvestor/the-time-has-come-for-a-new-type-of-database->

Arten von TSDBs¹²

Aufteilung von TSDBs in vier verschiedene Bereiche nach Bader, Kopp und Falkentahl

Verbreitung von TSDBs¹³¹⁴

Ranking von DB-Engines zur Verbreitung von TSDBs

Anwendung von TSDB

- Smart Metering^{15 16}

Erläuterungen zu Smart Metering Device und Funktion innerhalb des Smart Grid

Wissenschaftlicher Vergleich von Datenbanken im Kontext des Smart Meterings:¹⁷

„The conclusion is, that if the data and therefore the queries are based on a key like client identifier, sensor identifier, or timestamp, then some DBMS result in increased performance. The performance is logarithmically decreased for these DBMS when using a combination of tags.

Two of the compared DBMS are interesting for their scenario, depending whether the focus lies on INS queries, READ queries, or both. Hypertable is the best choice if the focus lies on the highest rate of executed INS queries in combination with a worse SCAN performance. BerkeleyDB, which has a lower rate of executed INS queries but executes more SCAN queries, is the second best choice.“

¹² vgl. Bader, A. B., Kopp, O. K., & Falkentahl, M. F. (2017). Survey and Comparison of Open Source Time Series Databases. Abgerufen 15. Juli, 2019, von ftp://ftp.informatik.uni-stuttgart.de/pub/library/ncstrl.ustuttgart_fi/INPROC-2017-06/INPROC-2017-06.pdf

¹³ vgl. Time Series Databases explained. [Bild/Grafik]. (o.D.). Abgerufen von <https://www.influxdata.com/time-series-database/>

¹⁴ DB-Engines Ranking of Time Series DBMS. (2019, Juli). Abgerufen von <https://db-engines.com/en/ranking/time+series+dbms/all>

¹⁵ Intelligente Stromzähler [Bild/Grafik]. (o.D.). Abgerufen von https://de.wikipedia.org/wiki/Intelligenter_Z%C3%A4hler#/media/Datei:Intelligenter_zähler-Smart_meter.jpg

¹⁶ Was sind eigentlich "Smart Grids"? (2019). Abgerufen 24. Juli, 2019, von <https://www.bmwi-energiende.de/EWD/Redaktion/Newsletter/2019/05/Meldung/direkt-erklaert.html>

¹⁷ Fortiș, T. F. Ciprian Pungilă, Teod, Aritoni, O. A., & Pungilă, C. P. (2014, 1. September). Benchmarking Database Systems for the Requirements of Sensor Readings. Abgerufen 24. Juli, 2019, von <https://www.tandfonline.com/action/cookieAbsent>

IoT - Internet of Things^{18 19 20} Erklärung zum Begriff Anwendungsbereich der TSDBs

- Monitoring Systems
Verschiedene Bereiche

Demo InfluxDB

Bedienung über die Shell:

- Erläuterungen zum CSV Datensatz
- Einpflegen des Datensatzes (Währungsverlauf von Bitcoins)
- Query Abfragen

Bedienung über den Browser

- Erklärungen zum Interface der Datenbank
 - Explore (Testqueries)
 - Dashboards (zur Abspeicherung von Queries)
- Erstellung eines Monitoring Tools zur Systemauslastung des Rechners
 - Grafische Auswertung
 - Tabellarische Auswertung
- Erstellung eines Dashboards für Monitoring

Fazit

Zeitreihenbanken sind gerade in der Entstehung und werden als die nächste große Entwicklung auf den Markt der Datenbanken angesehen. Sie eignen sich ideal für das Verwalten und Bearbeiten von großen Datenmengen und werden zukünftig in einer immer weiter vernetzten Welt eine große Rolle spielen.

Man spricht auch nicht von „dem“ Zeitreihendatenbanksystem, sondern es gibt eine Vielzahl unterschiedlicher Architekturen von Datenbanken die den Umgang mit

¹⁸ Luber, S. L. (2016, 1. September). Was ist das Internet of Things? Abgerufen 24. Juli, 2019, von <https://www.bigdata-insider.de/was-ist-das-internet-of-things-a-590806/>

¹⁹ Whitlock, S. W. (2019, 25. März). 10 Principles of the Internet of Things (IoT) | | Flexware Innovation [Bild/Grafik]. Abgerufen 24. Juli, 2019, von

<https://www.flexwareinnovation.com/10-principles-of-the-internet-of-things-iot/>

²⁰ Internet of Things Definition. (o.D.). Abgerufen 24. Juli, 2019, von https://www.gruenderszene.de/lexikon/begriffe/internet-of-things?interstitial%3Finterstitial?interstitial_click

Zeitreihen beherrschen. Bei der Verwendung solch einer Datenbank ist es wichtig Anforderungen im Vorfeld genau zu kennen und zu definieren, um entscheiden zu können welche Art von Zeitreihendatenbanken die Richtige für die jeweilige Anwendung ist.

Quellenangaben:

- vgl. Einführung in die Zeitreihenanalyse, Kreiß, J. P. K., & Neuhaus, G. N. (o.D.). Google Books. Abgerufen 15. Juli, 2019, von <https://books.google.de/?hl=de>
- vgl. Bader, A. B., Kopp, O. K., & Falkentahl, M. F. (2017). Survey and Comparison of Open Source Time Series Databases. Abgerufen 15. Juli, 2019, von ftp://ftp.informatik.uni-stuttgart.de/pub/library/ncstrl.ustuttgart_fi/INPROC-2017-06/INPROC-2017-06.pdf
- vgl. Fangman, S. F. (2019, 28. Januar). Data Driven Investor. Abgerufen von <https://medium.com/datadriveninvestor/the-time-has-come-for-a-new-type-of-database-47cf8df1667a>
- vgl. Time Series Databases explained. (o.D.). Abgerufen von <https://www.influxdata.com/time-series-database/>
- Key/Value-Datenbanksysteme. (o.D.). Abgerufen von <http://www.datenbanken-verstehen.de/lexikon/key-value-datenbanksysteme/>
- Time Series Databases: New Ways to Store and Access Data. (2014). Abgerufen von <https://apprize.info/data/series/4.html>
- vgl. InfluxDB key concepts. (o.D.). Abgerufen von https://docs.influxdata.com/influxdb/v1.7/concepts/key_concepts/
- vgl. Time Series Databases explained. [Bild/Grafik]. (o.D.). Abgerufen von <https://www.influxdata.com/time-series-database/>
- DB-Engines Ranking of Time Series DBMS. (2019, Juli). Abgerufen von <https://db-engines.com/en/ranking/time+series+dbms/all>
- Intelligente Stromzähler [Bild/Grafik]. (o.D.). Abgerufen von https://de.wikipedia.org/wiki/Intelligenter_Z%C3%A4hler#/media/Datei:Intelligenter_zaeher-Smart_meter.jpg
- Was sind eigentlich "Smart Grids"? (2019). Abgerufen 24. Juli, 2019, von <https://www.bmwi-energiewende.de/EWD/Redaktion/Newsletter/2019/05/Meldung/direkt-erklaert.html>

- Fortiș, T. F. Ciprian Pungilă, Teod, Aritoni, O. A., & Pungilă, C. P. (2014, 1. September). Benchmarking Database Systems for the Requirements of Sensor Readings. Abgerufen 24. Juli, 2019, von <https://www.tandfonline.com/action/cookieAbsent>
- Luber, S. L. (2016, 1. September). Was ist das Internet of Things? Abgerufen 24. Juli, 2019, von <https://www.bigdata-insider.de/was-ist-das-internet-of-things-a-590806/>
- Whitlock, S. W. (2019, 25. März). 10 Principles of the Internet of Things (IoT) | Flexware Innovation [Bild/Grafik]. Abgerufen 24. Juli, 2019, von <https://www.flexwareinnovation.com/10-principles-of-the-internet-of-things-iot/>
- Internet of Things Definition. (o.D.). Abgerufen 24. Juli, 2019, von https://www.gruenderszene.de/lexikon/begriffe/internet-of-things?interstitial%3Finterstitial?interstitial_click