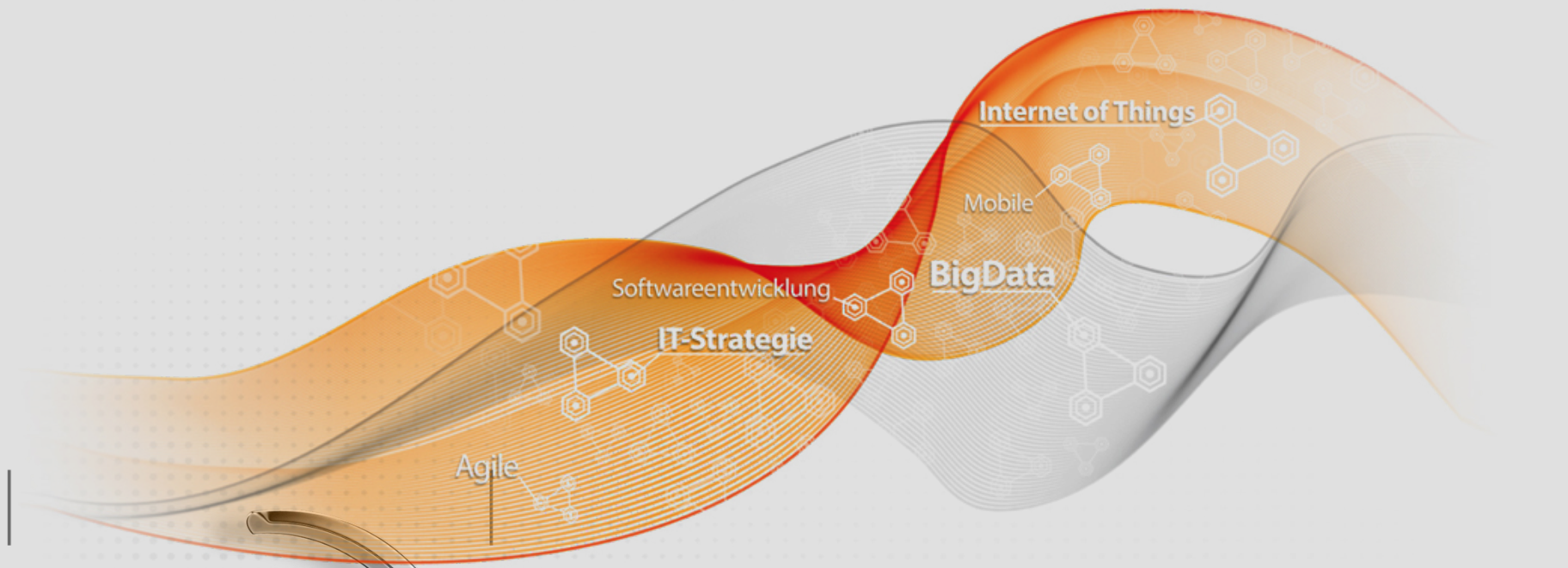




IT Consulting

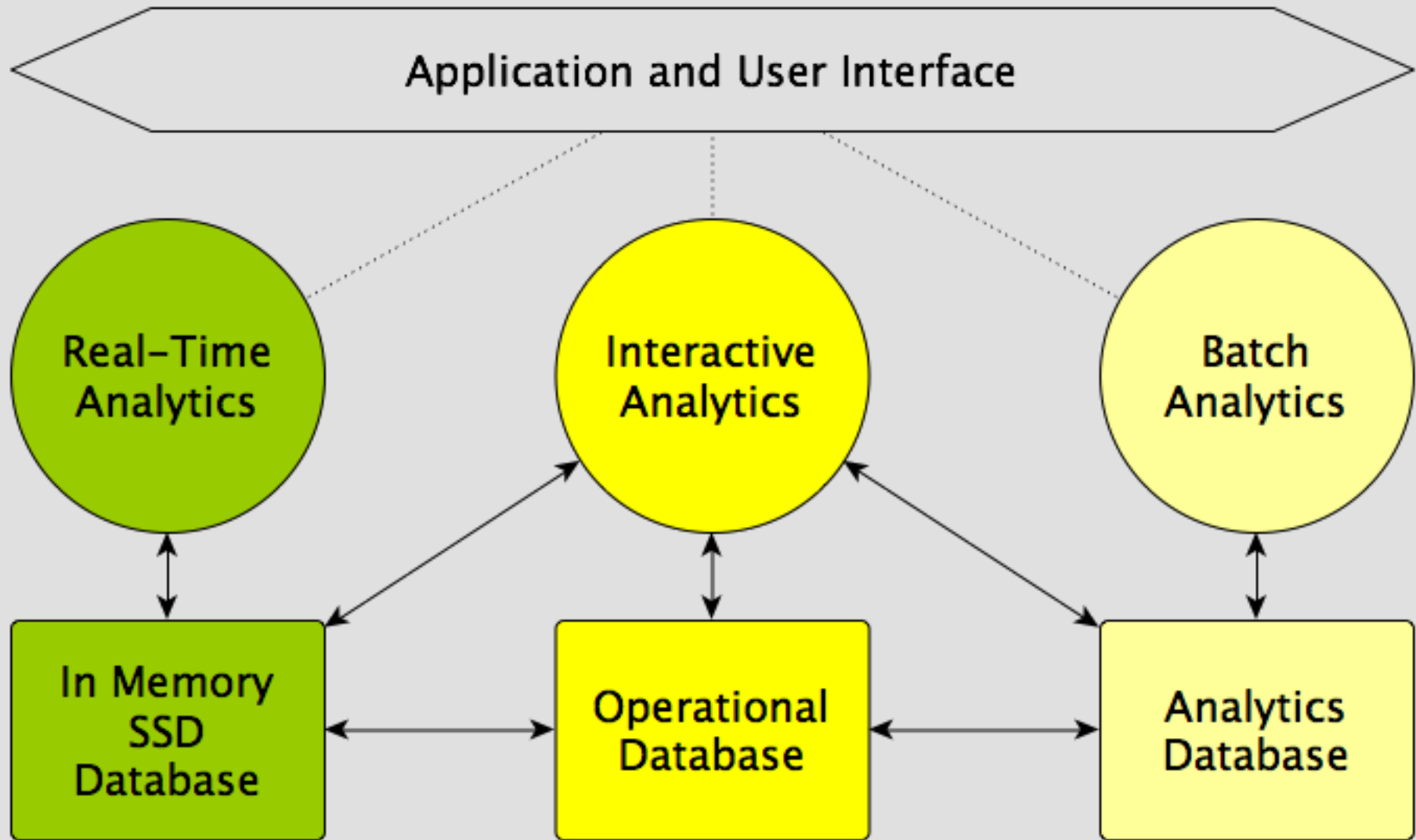
Real-Time Analytics mit Spark und Cassandra

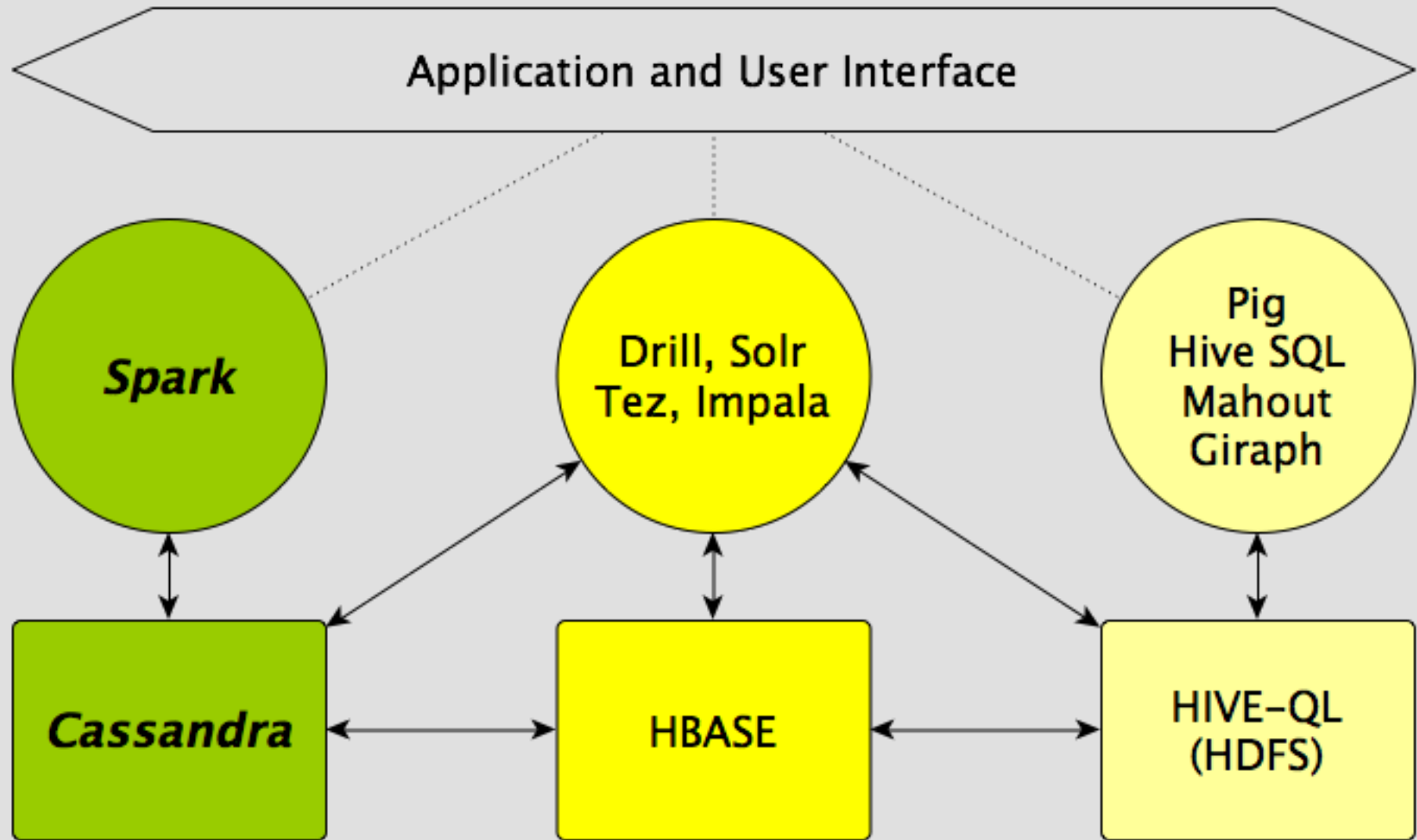
im März 2015
OSBI – Workshop
<http://www.osbi-workshop.de/>



REAL-TIME ANALYTICS?









CASSANDRA



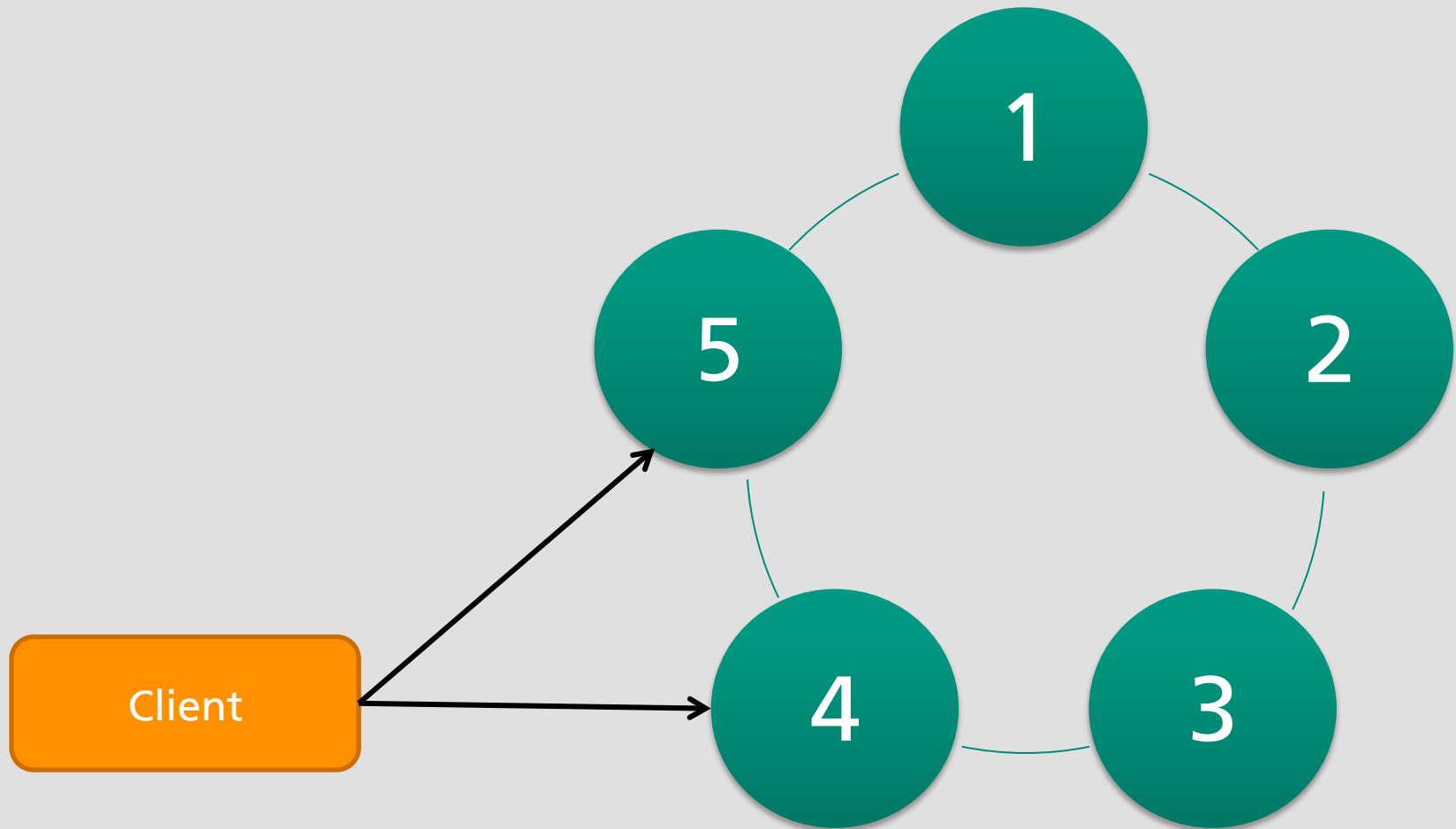


Vorteile Cassandra: massiv skalierbare verteilte Datenbank

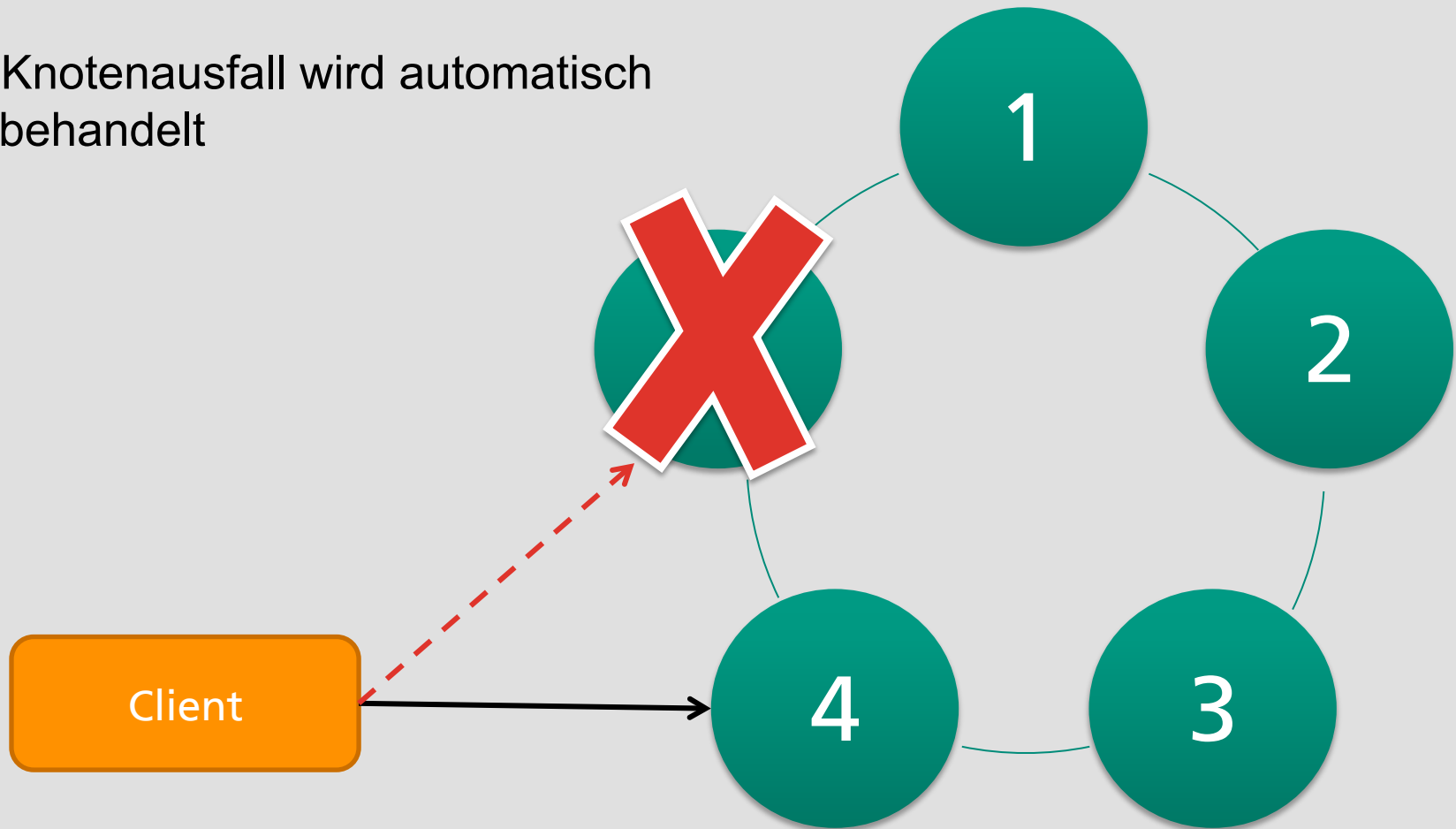
- 1** CAP frei einstellbar, für Analytics: AP
- 2** Shared Nothing, Peer to Peer
- 3** KKV – Wide Columnar/Partitions
- 4** Zeitreihen optimierte Datenmodelle
- 5** In Memory Tabellen
- 6** Daten-Lokalität mit Wide Partitions



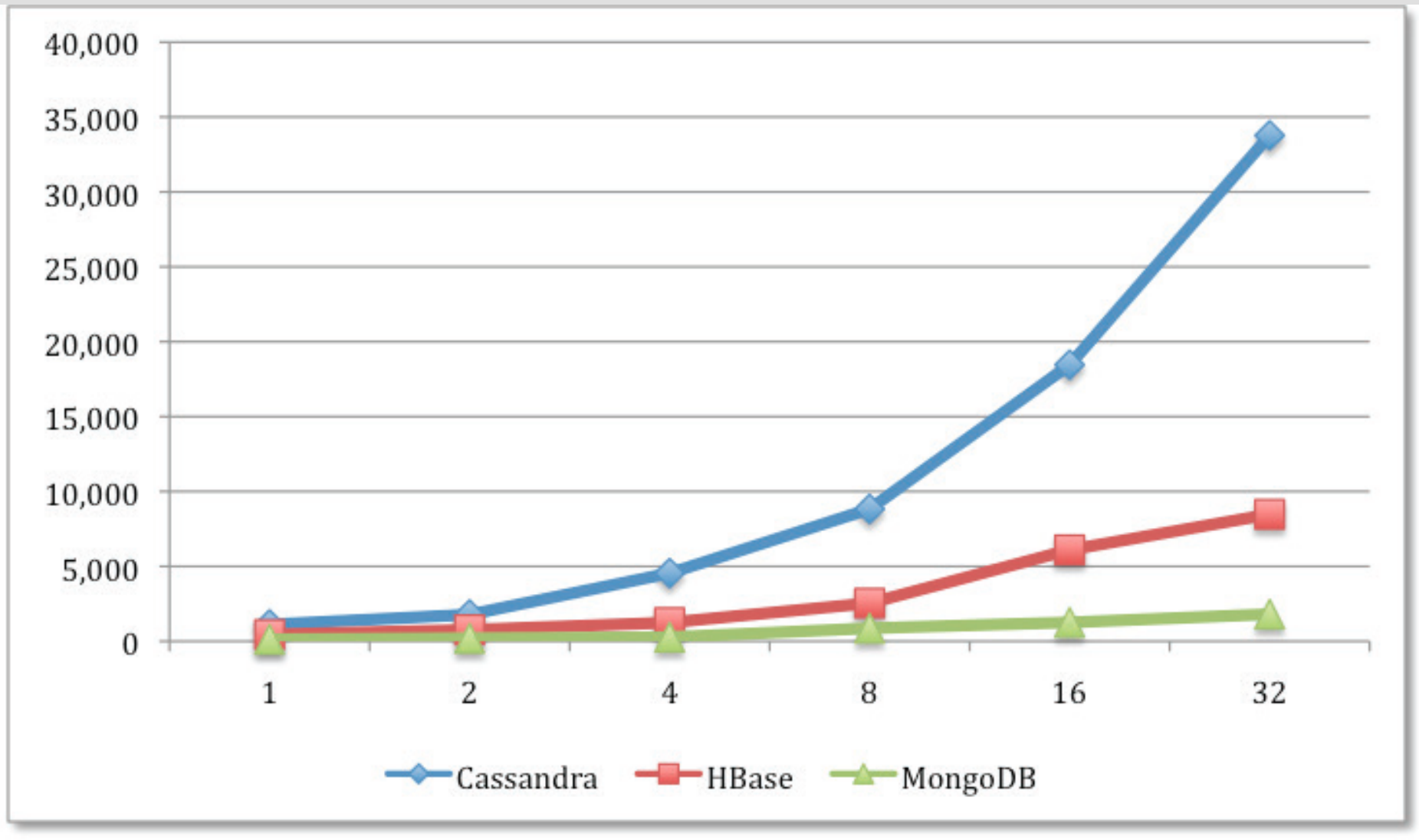
Jeder Knoten in Cassandra ist äquivalent ansprechbar
Konfigurierbare Replikation (lokal, DC-weise)



- „Can't Fail, Must Scale“ –System
- Datenreplikation sichert Verfügbarkeit
- Knotenausfall wird automatisch behandelt



Anzahl Operationen (Read/Write) pro Sec – Anzahl Knoten



Quelle: Planet Cassandra



SPARK



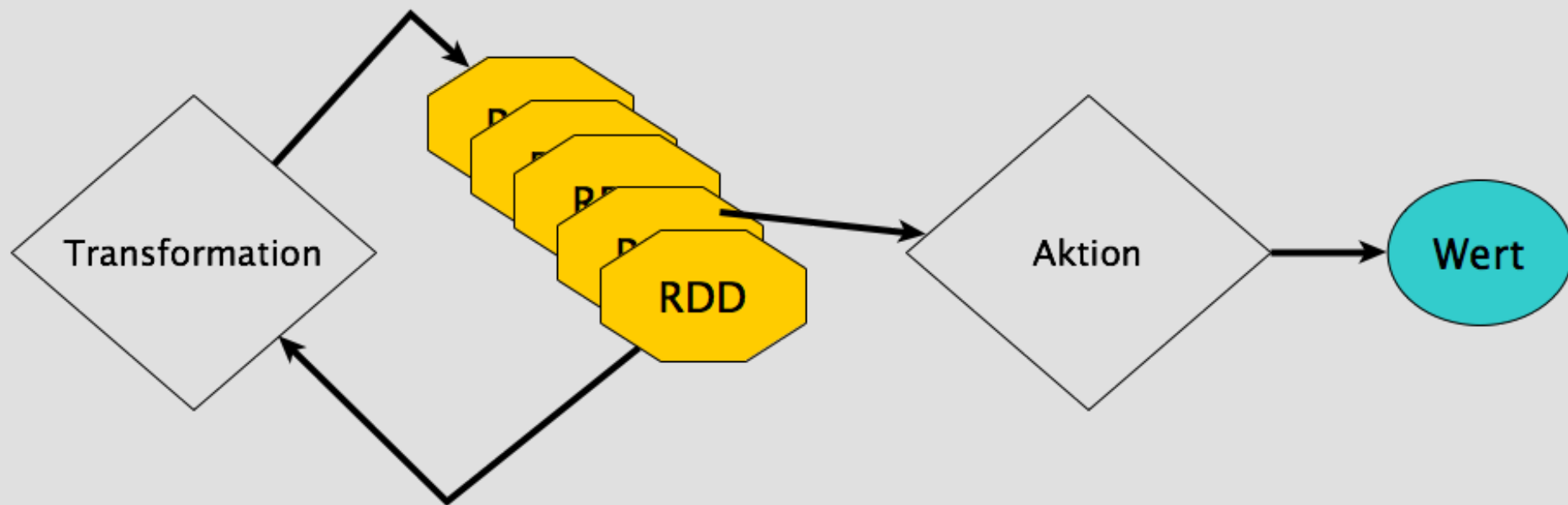
Kernelemente





Resilient Distributed Datasets





Transformationen

map
filter
groupByKey
join
...

Aktionen

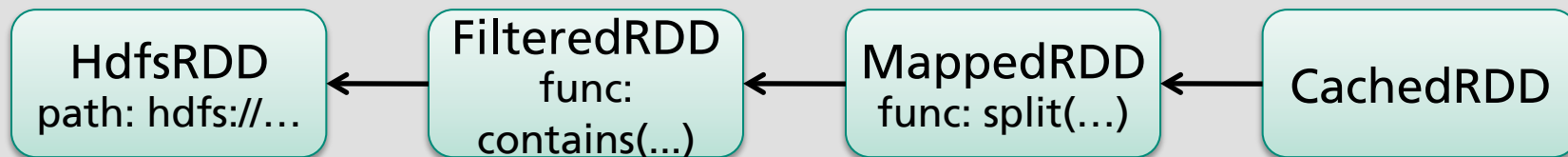
reduce
collect
count
lookupKey
...



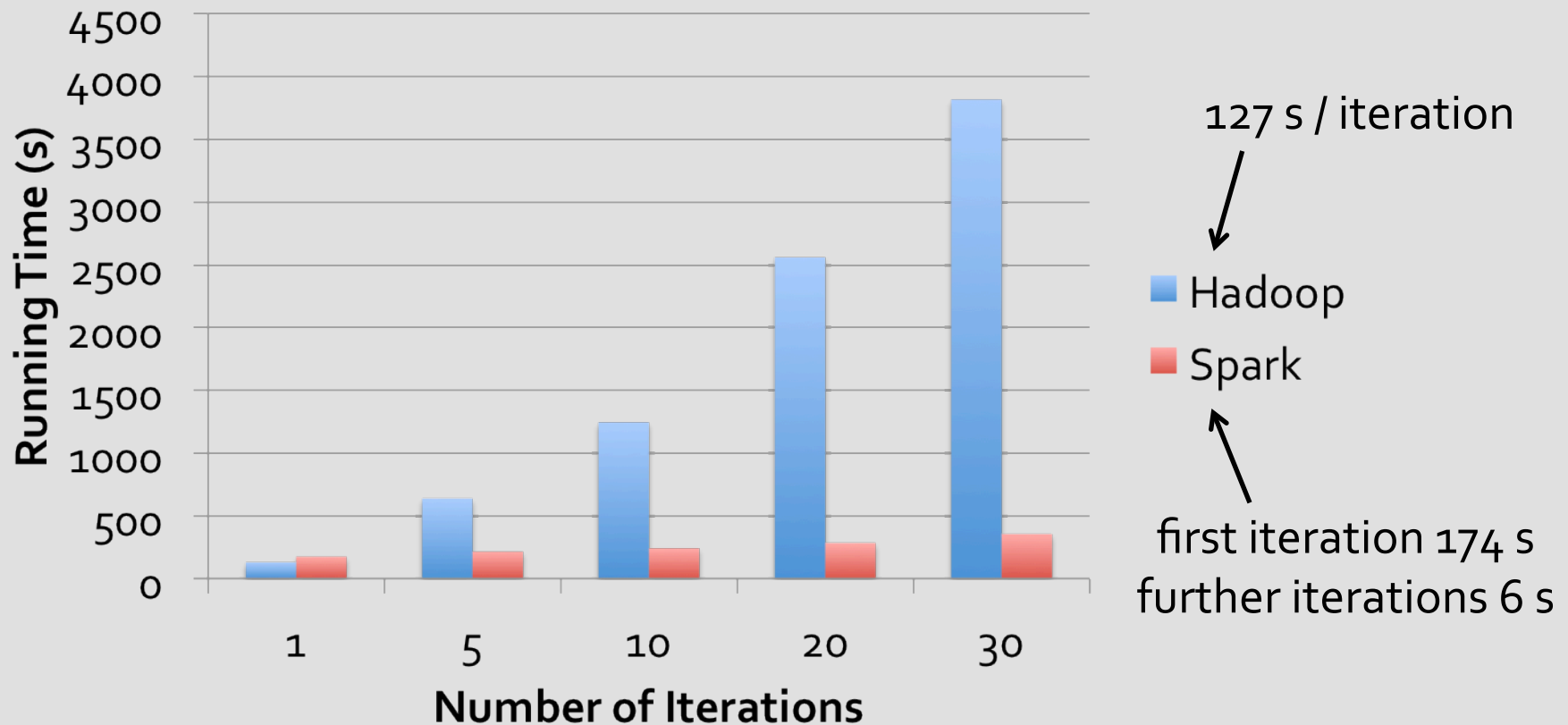
Resilient Distributed Datasets

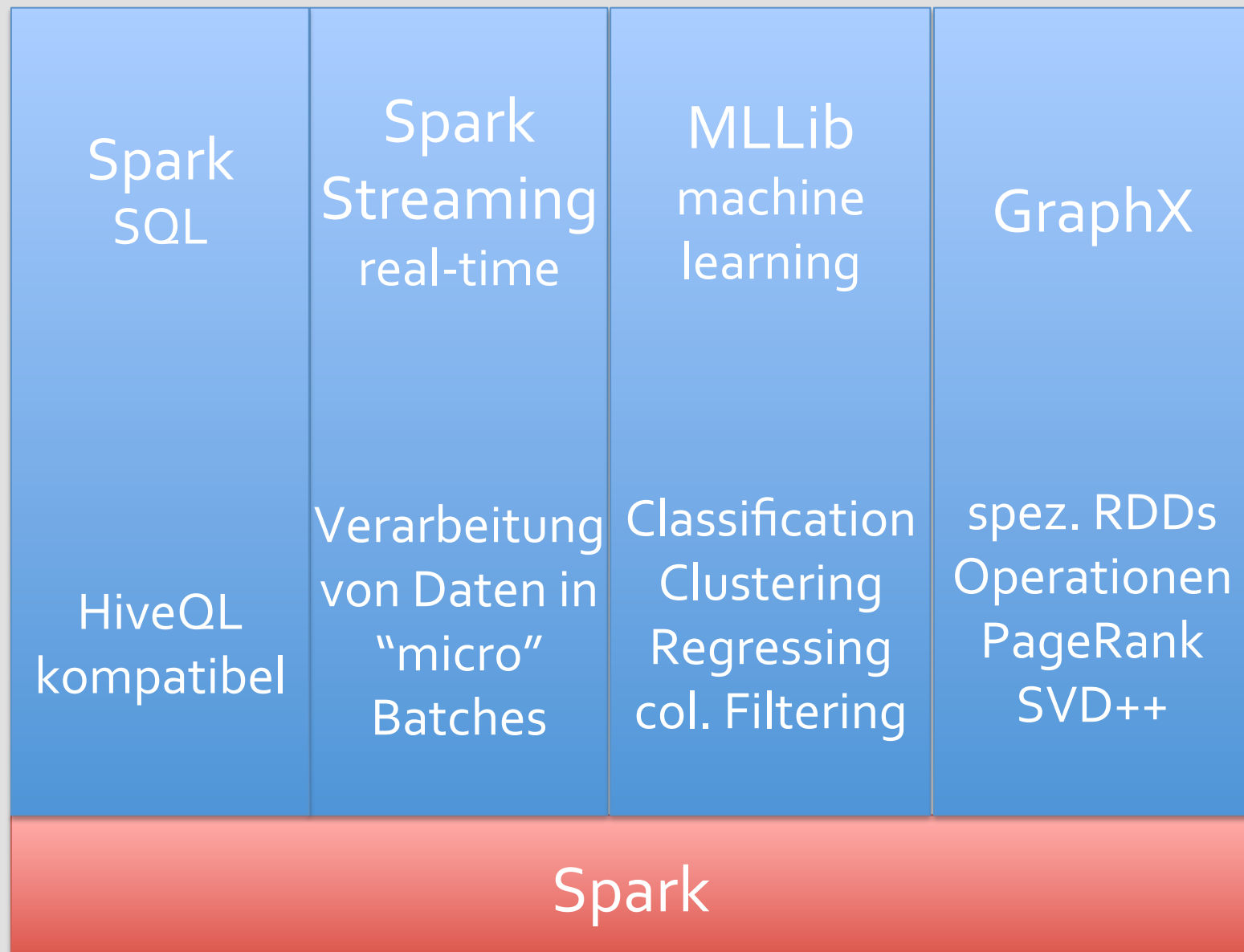
speichern ihre Herkunft (Lineage)

damit kann jederzeit bei Ausfall
die entsprechende Partition der RDD neu aufgebaut werden



Logistic Regression





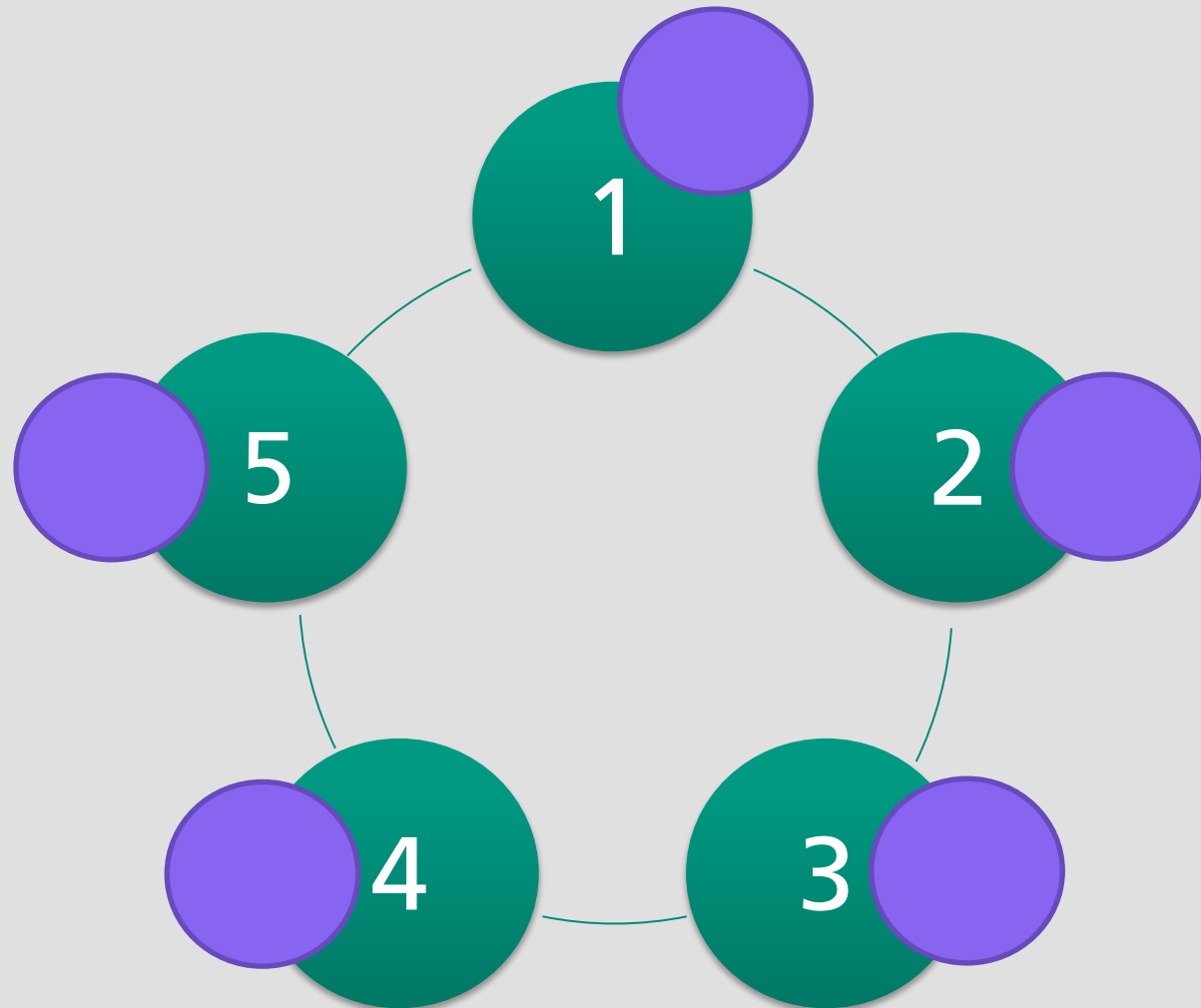


Integration

Spark & Cassandra



mit Spark-Cassandra Connector
Cassandra Tabellen sind als RDDs verfügbar
auf jedem Cassandra Knoten wird ein Spark Executor eingesetzt

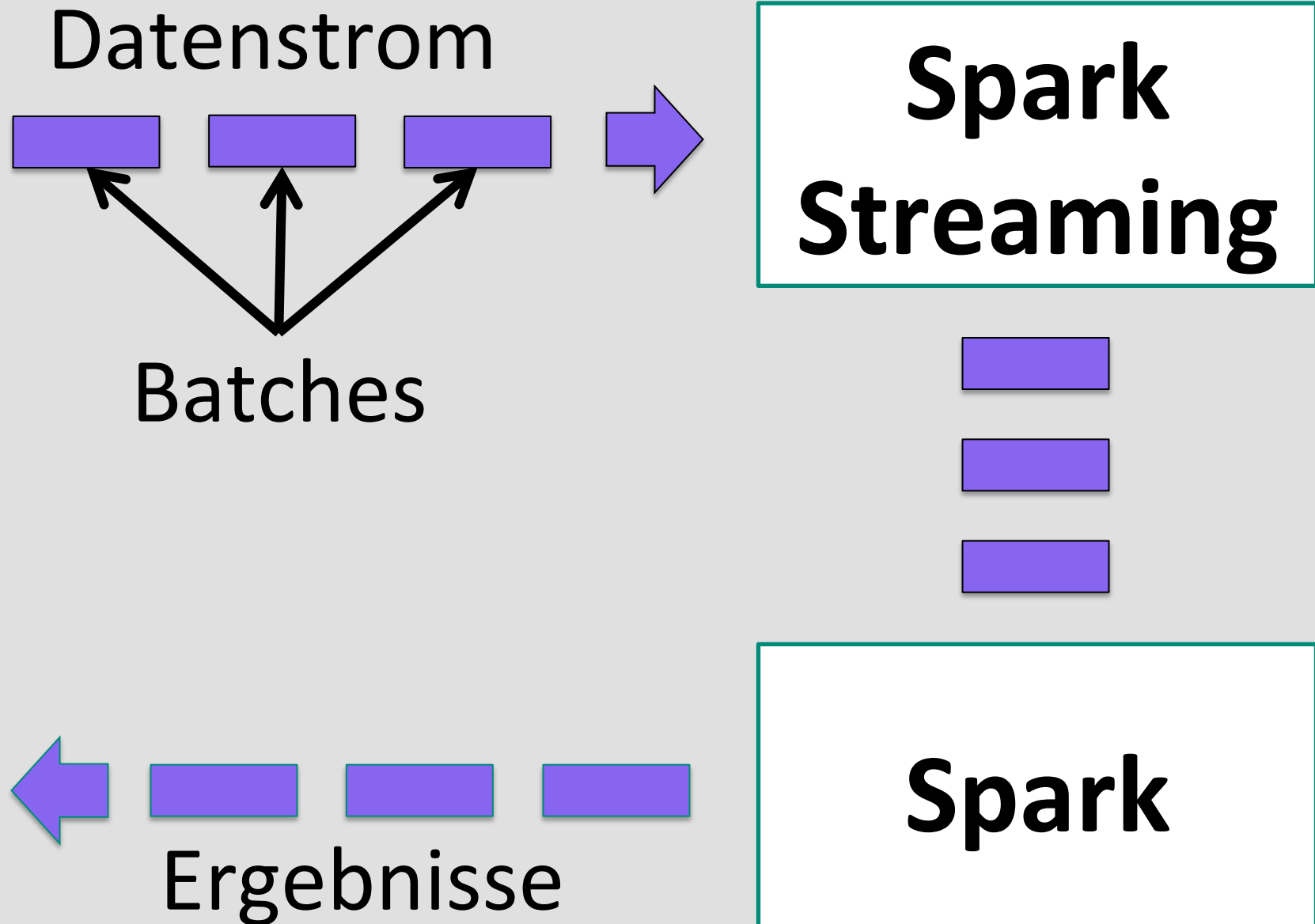






Spark Streaming





**1**

Integrierbar mit Cassandra/Spark Treiber

2

Micro Batches (1 Sek), Discretized Streams

3

Exactly Once Semantik

4

RDD Funktionalität

5

Integration diverser MQ (z.B. Kafka)



Umfeld

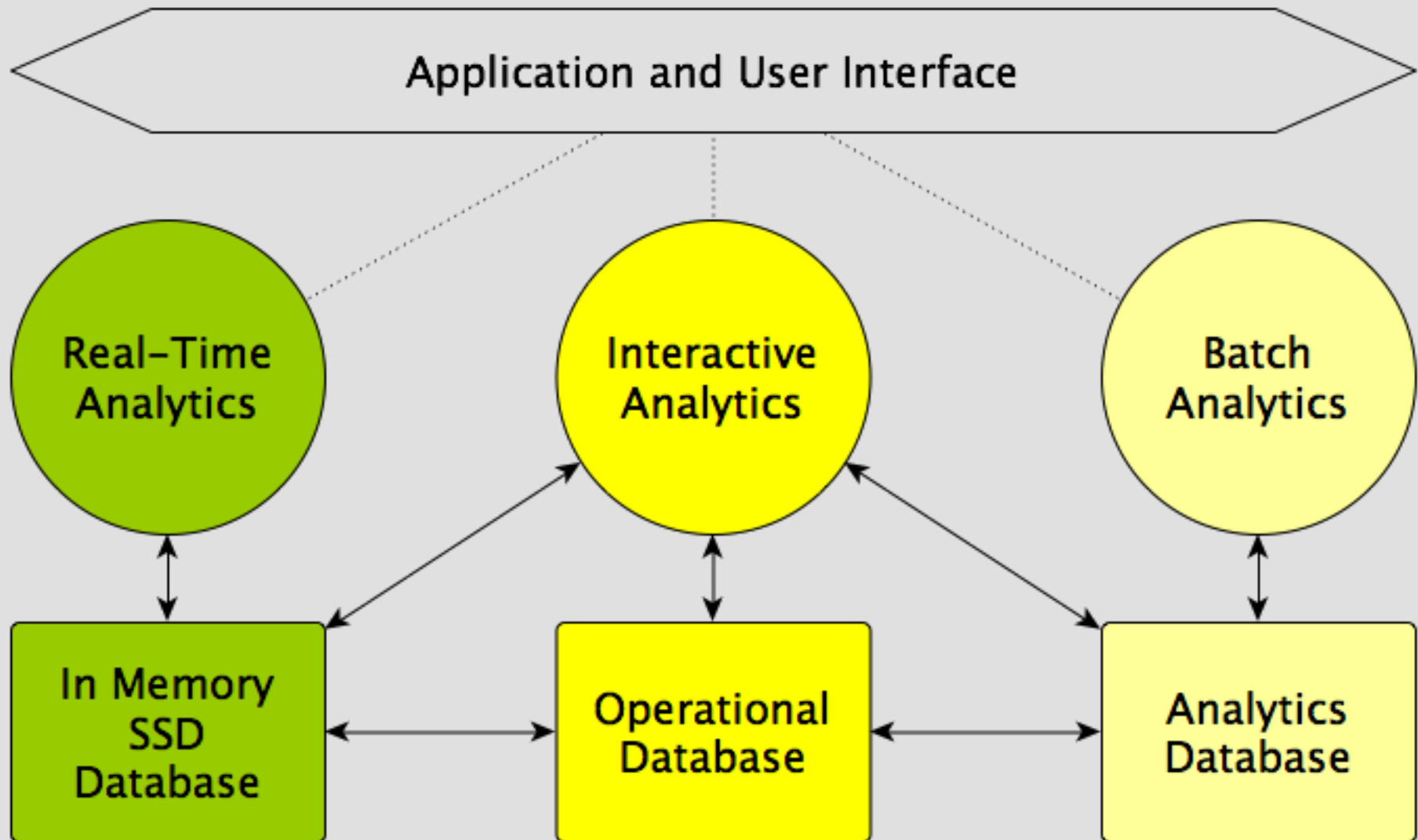


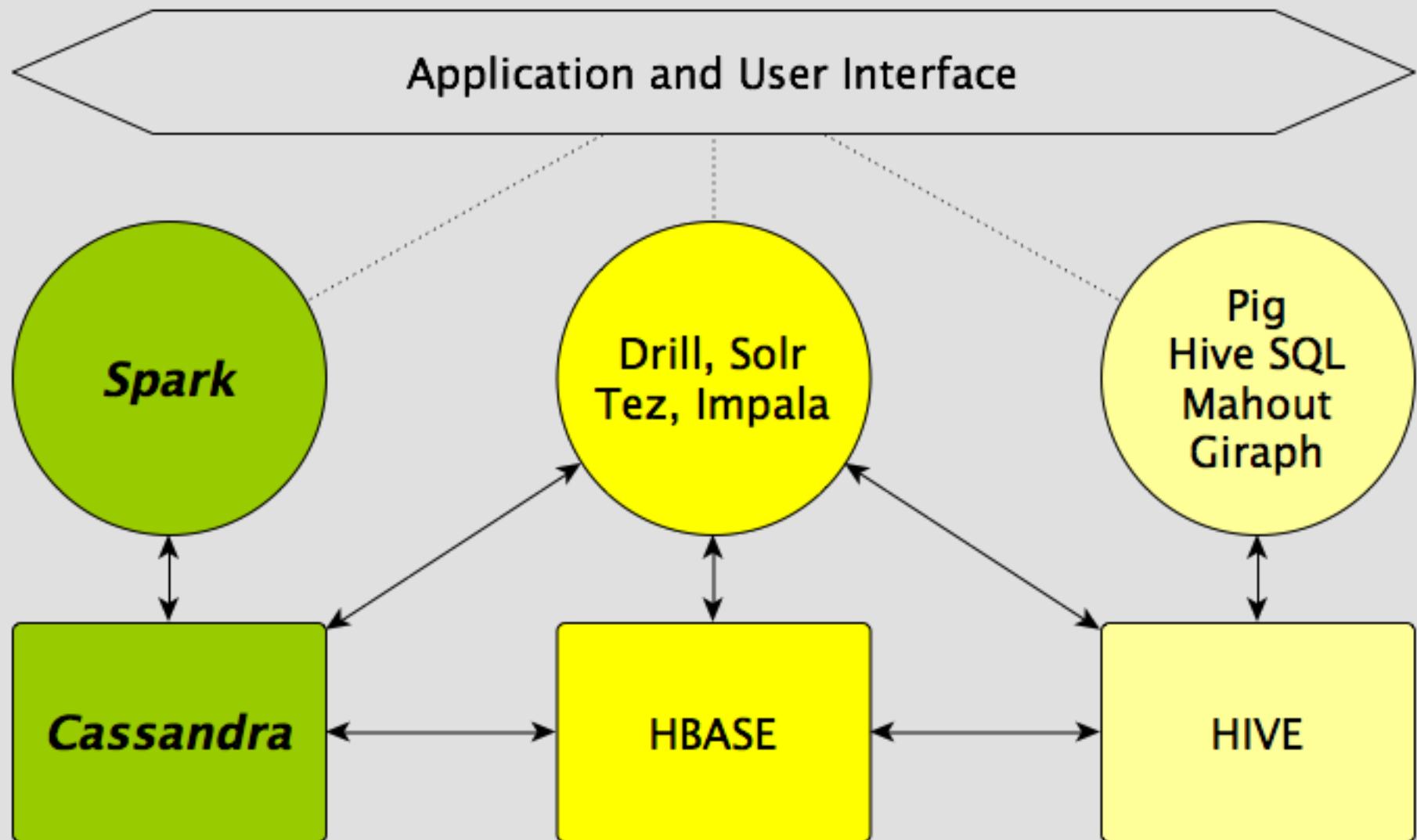


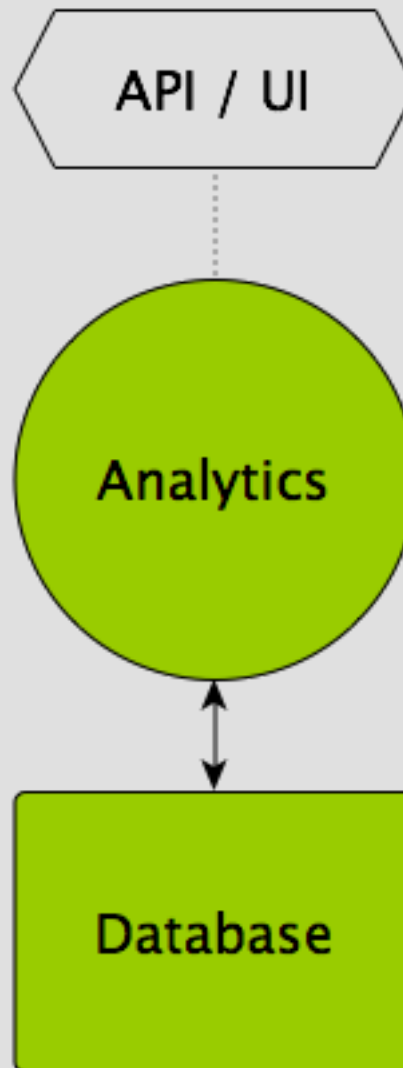


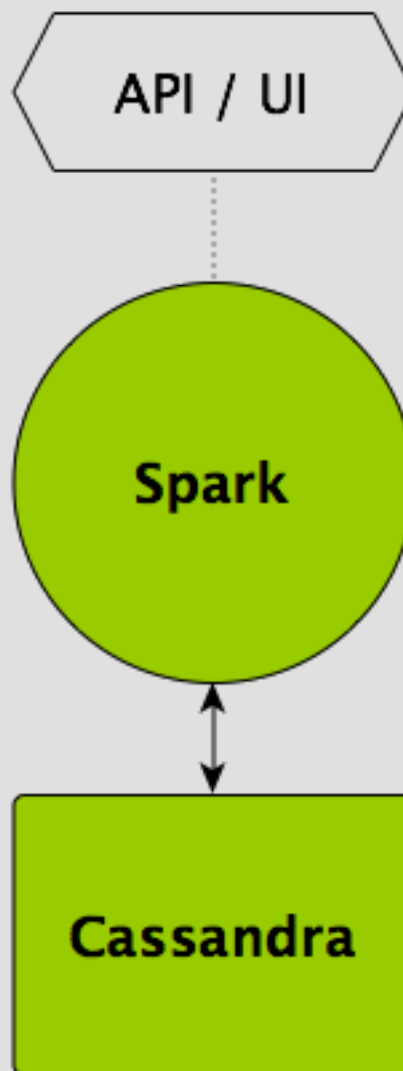
In Zukunft ?











Wir beraten Sie gerne

HighEnd-Technology requires HighEnd-Competence

WidasConcepts

Thomas Mann, Solution Architect

Telefon: +49 (7044) 95103 – 100

Mobile: +49 162 259 56 90

Mail: thomas.mann@widas.de

WidasConcepts GmbH

Maybachstraße 2
71299 Wimsheim

www.widas.de



Dieses Dokument wurde von WidasConcepts erstellt. Die Verteilung, Zitierung und Vervielfältigung – auch auszugsweise – zum Zwecke der Weitergabe an Dritte ist nur mit vorheriger schriftlicher Zustimmung von WidasConcepts gestattet.

This presentation was created by WidasConcepts. Distribution, citation, copying - completely or in extracts – for transfer purposes, is only permitted with prior written agreement. These abstracts and graphics were deployed by WidasConcepts within the scope of a presentation. It is no complete documentation of this event.