

EXASOL @ Symposium on Scalable Analytics

www.exasol.com

Skalierbare Analysen mit EXASolution

EXASOL AG - Wer sind wir

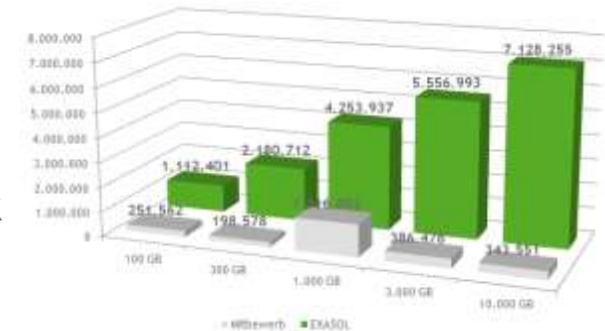
R&D: + seit 2000
+ laufend Forschungsprojekte

Produkt: Analytische Datenbank EXASolution
Focus auf Komplexität in der Auswertung &
High-Performance-Szenarien

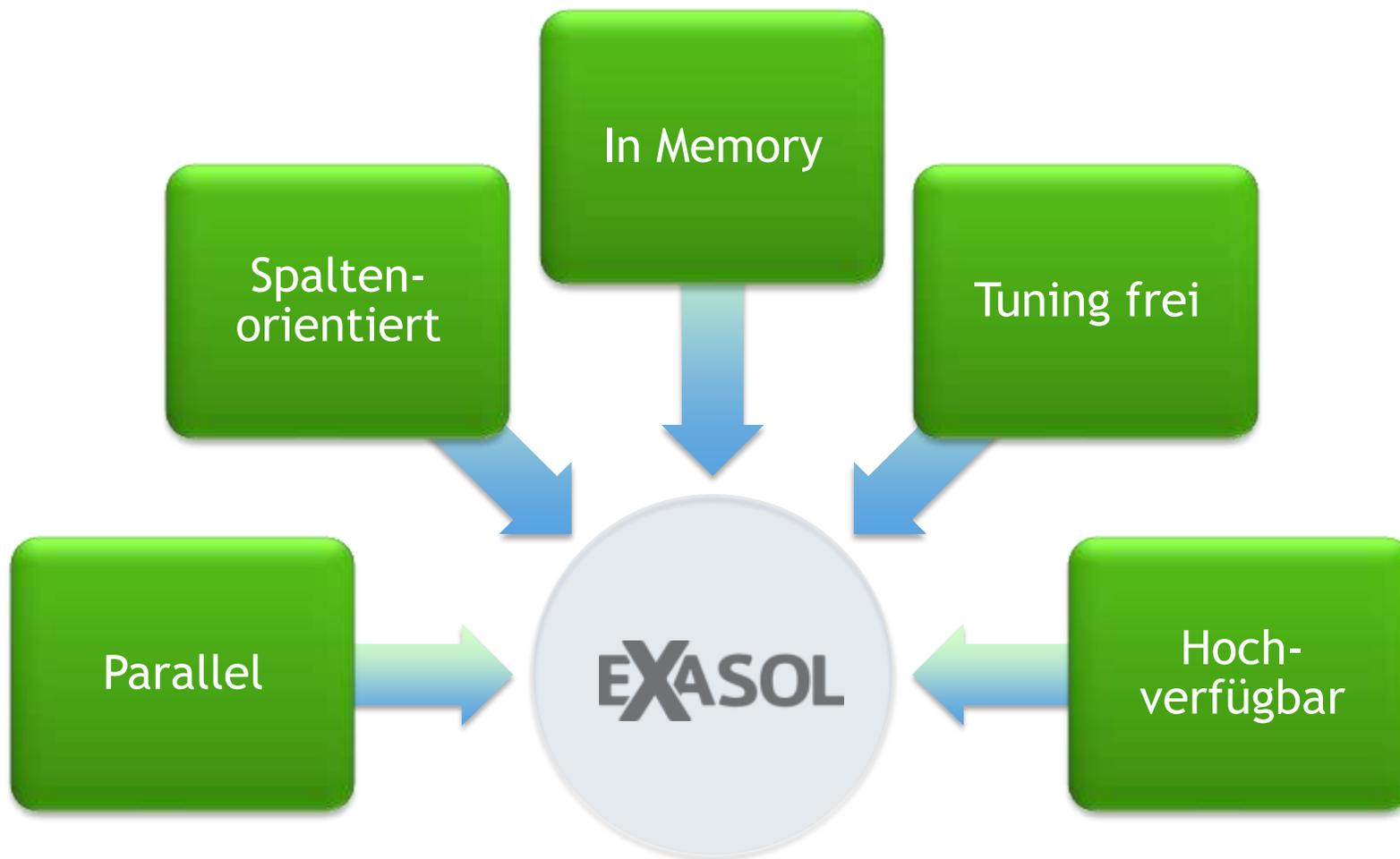
Erweiterungen: EXAPowerlytics (R, LUA, Python)
+ In-Database-Analytics

Firmensitz: Nürnberg, Deutschland

Key Facts: + Skalierbarkeit über Standard-Server
+ Performance-Führerschaft im TPC-H Benchmark
+ Seit 2010 erfolgreichster Anbieter analytischer
Datenbanken in Deutschland (BARC)



EXASolution: Architekturüberblick



EXASolution Under the Hood



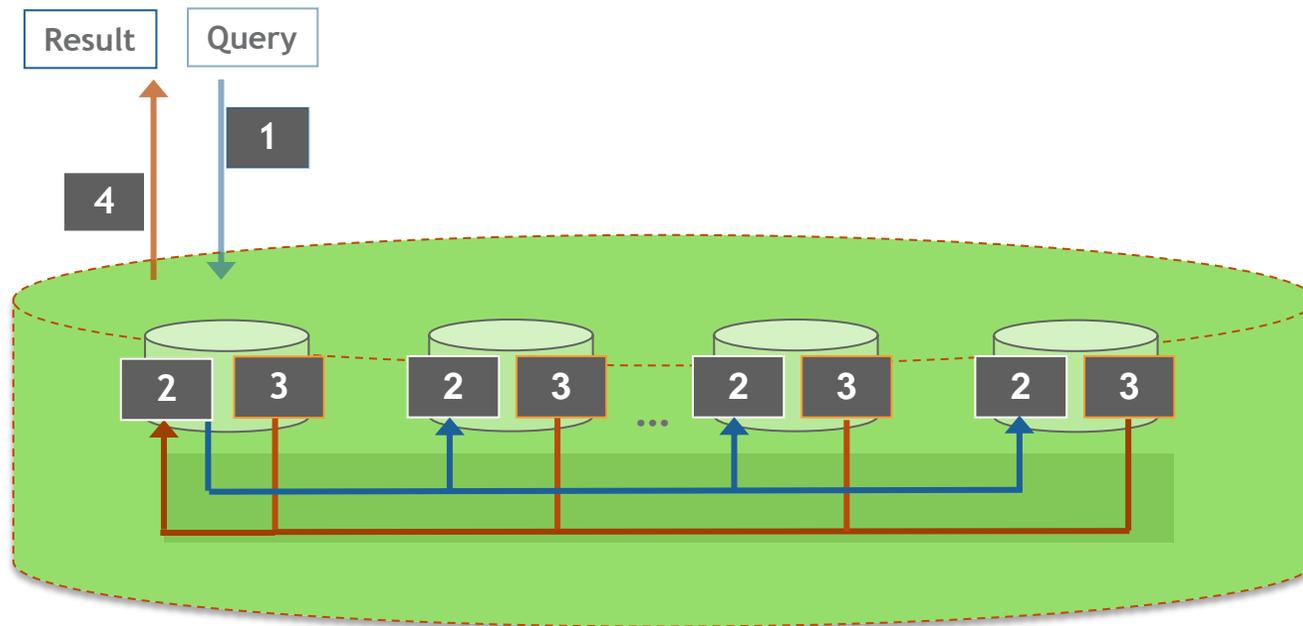
- ✓ Shared nothing architecture, SPMD paradigm
- ✓ Developed from scratch - whole design focused on clustering
 - Competing systems usually built on top of existing open source kernels
- ✓ Designed for In Memory processing, not just a big cache

- ✓ Data is persistently stored on local disks
- ✓ Vertical / Column-based storage
- ✓ Special compression techniques (high performance needs)

- ✓ Full of automated processes
 - ✓ Automatic data distribution and replication
 - ✓ Automatic & transparent creation of highly compressed join indices
 - ✓ Automatic compression method selection for maximum ratios

Abfrage-Performance

Cluster Shared-Nothing-Architektur für massiv parallele Datenverarbeitung



Abfrage-Performance

- Spaltenorientierte Datenspeicherung
- Sehr effiziente Kompression



Kunden					
	ID	Name	Vorname	Umsatz	...
Row	1	Becker	Hans	23000	
	3	Weber	Peter	730000	
	4	Huber	Klaus	39600	
	5	Schmidt	Maria	124000	
	6	Schneider	Thomas	93600	
	22	Fischer	Stefan	368200	

AVG

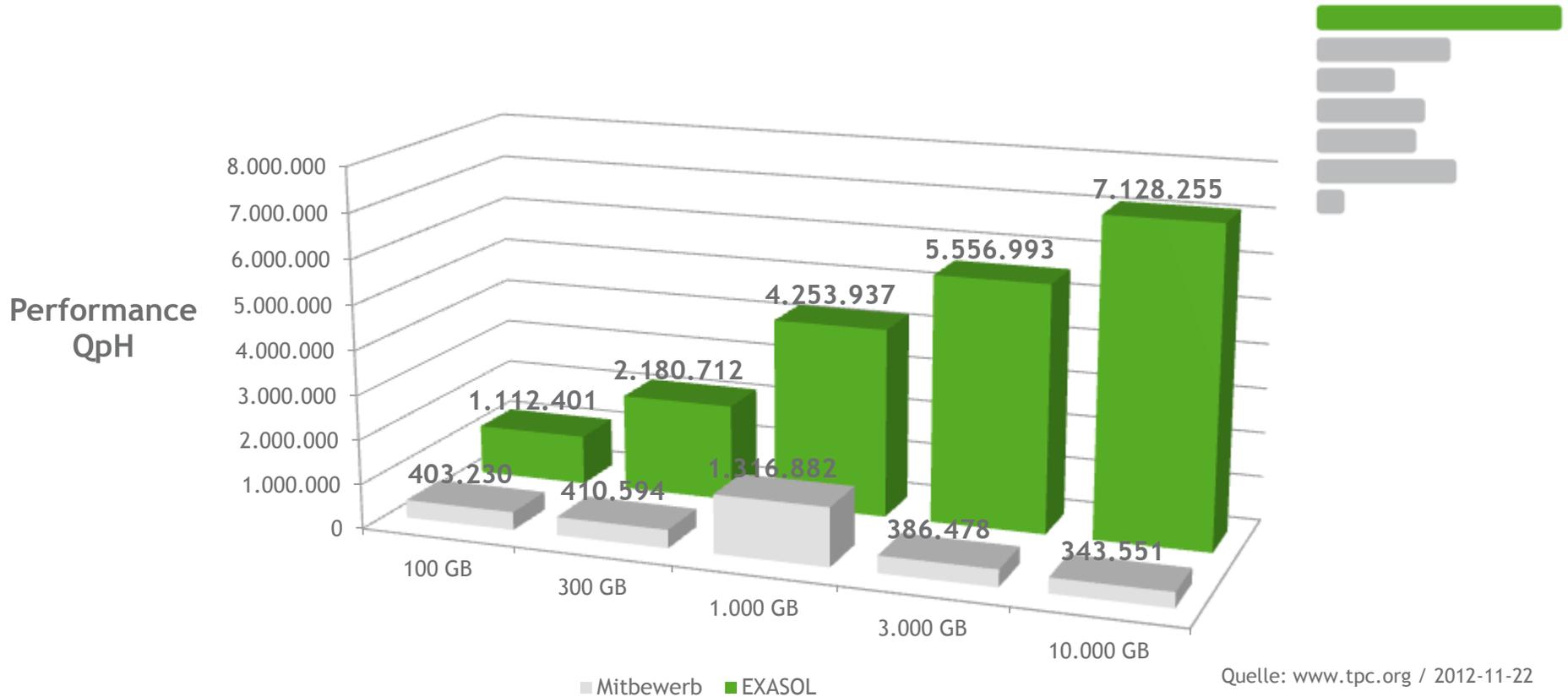
Abfrage-Performance

In-Memory Technologie

- Hybrid-Ansatz (In-Memory-Fokussierung mit Festplattenunterstützung)
- Priorisierung wichtiger Daten
- Intelligente Algorithmen für Hauptspeicher - Verarbeitung
- Random-Zugriff auf einzelne Werte
- Daten werden 1:1 in Hauptspeicher „gemappt“
- Keine Konvertierung der Datenstrukturen (Dekompression)
- Optimierte Least recently used (LRU) - Mechanismen

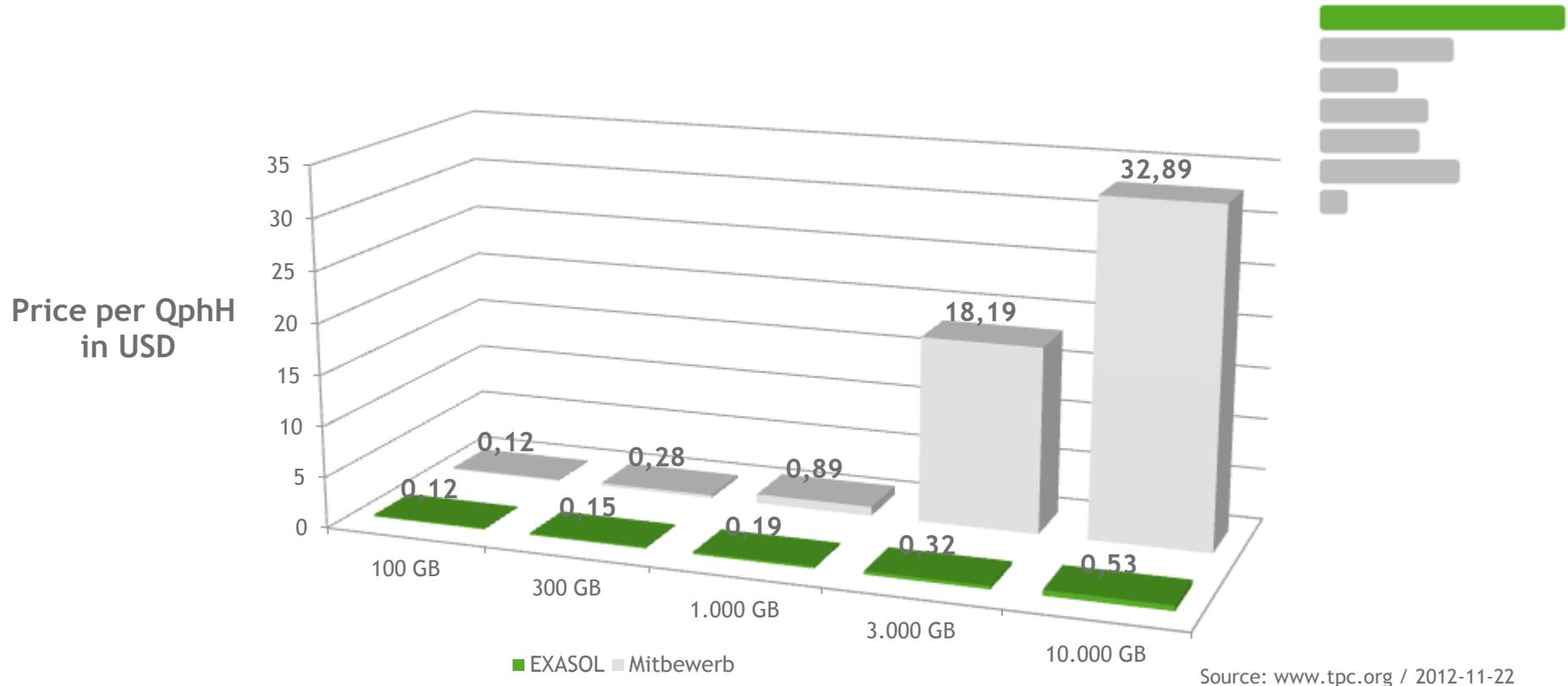


Alle TPC-H 2011 Ergebnisse im Vergleich zum Wettbewerb



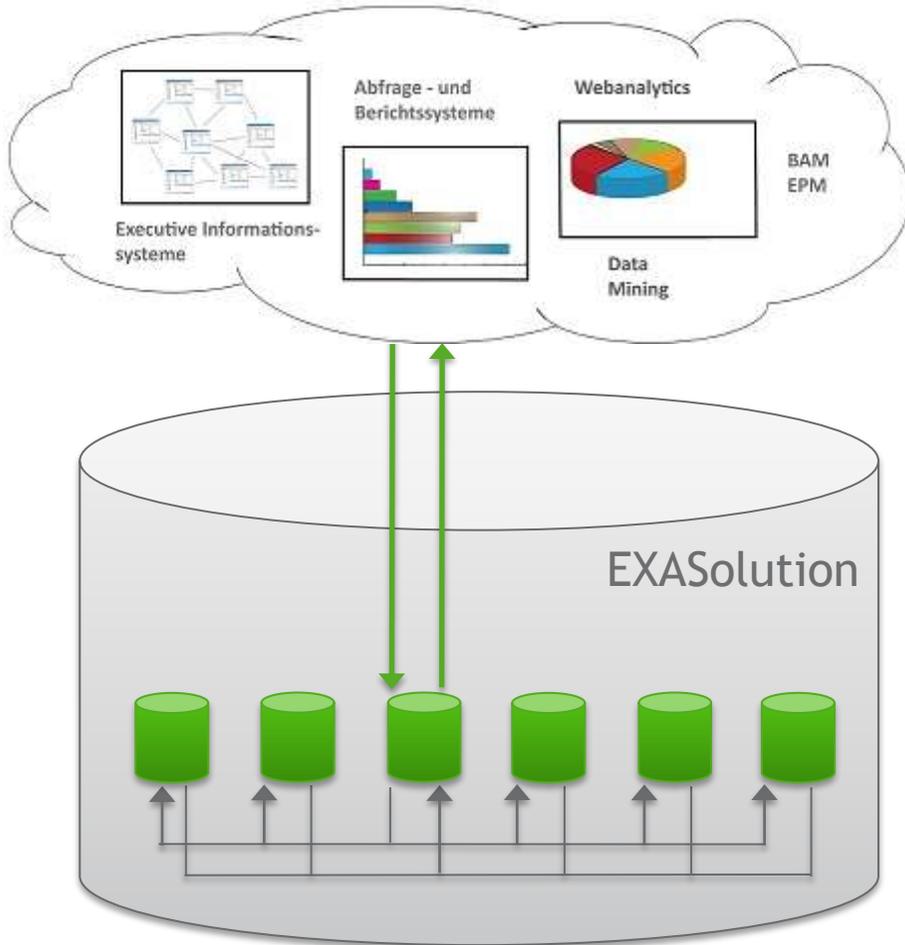
„Es ist bemerkenswert, dass EXASOL die für Data-Warehouse-Anwendungen besten Resultate in allen fünf Kategorien, von 100 GB bis hin zu 10 TB, erreicht hat“ (Dr. Carsten Bange, BARC)

Höchste Leistung bei jeweils bestem Preis-Leistungsverhältnis



„Zudem adressiert EXASOL mit seiner guten Kosten-Performance einen wunden Punkt vieler derzeit auf dem Markt befindlicher Systeme.“ (Dr. Carsten Bange, BARC)

Technischer Überblick



Standard Server

- RAM: 16 - 384 GB
- CPU: 2 Intel Xeon
- HDD: 2 - 24 SAS/SATA
- Ethernet: Gbit / 10Gbit / Infiniband

Cluster "powered by" EXACluster OS

- Basiert auf Linux
- Einfache Installation und Administration
- Hochverfügbarkeit mit automatischer Ausfallsicherung

EXASolution: RDBMS für komplexe Analysen

- Massiv parallel
- In Memory
- Spaltenorientiert
- Tuning-frei
- Hochverfügbar

Reibungslose Integration

- Standard Schnittstelle(ODBC, JDBC, ADO.NET)
- Native CLI
- Hochperformanter Loader für DBMS, CSV, FBV
- SQL Editor EXAplus

Fragen? Vielen Dank für die Aufmerksamkeit

Wir suchen Verstärkung:

- + Software Entwickler Distributed Operating System (m/w)
- + Software Entwickler High Performance Cluster Computing (m/w)
- + C++ Software Entwickler (m/w)
- + Technical Consultant (SQL, Java) (m/w)

Kontakt:

www.exasol.com

Johannes Meier
johannes.meier@exasol.com

Oleksandr Kozachuk
oleksandr.kozachuk@exasol.com

