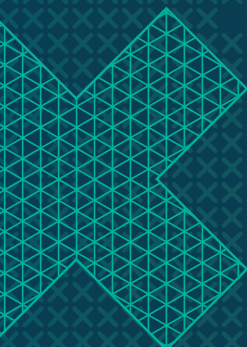




The world's fastest analytic database



Contents

The world's fastest analytic database

01

Introduction **3**

Key capabilities 4

02

What is Exasol? **6**

Core architecture 6

Other features 9

03

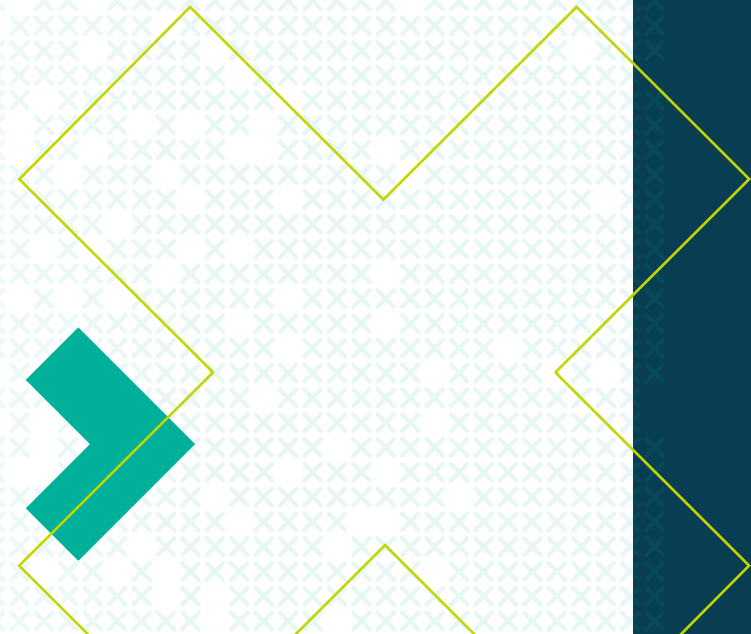
Value proposition **10**

04

Deployment benefits **12**

05

Conclusion **14**



Introduction

Exasol was founded in Nuremberg, Germany with a single purpose – to engineer the world’s fastest database for analytics, with no limits on data volumes. For more than a decade, the company has focused exclusively on delivering ultra-fast, massively scalable, analytic performance.

The company’s flagship product, Exasol, is a high-performance, in-memory, MPP database designed specifically for analytics. In 2011, Exasol set a new performance record in the TPC-H benchmark for clustered, decision support databases, and in 2014 extended these performance and price/performance records. Exasol remains in the number one position for all volumes of data, from 100GB right up to 100TB.

Supported from Exasol offices in Germany, UK, USA and Brazil and with partners across Europe, Israel and Japan, over 300 organisations are improving their business operational efficiencies and/or delivering excellent customer service using Exasol. Exasol customers span many market sectors including Digital Media, Retail, Communications, Financial Services, Manufacturing and Research.

Designed from the ground up, using state-of-the-art software techniques and principles, Exasol runs on low cost, commodity x86 hardware; scales from 10’s GBs to 100’s TBs data; is quick to implement and delivers extreme performance, without cost and complexity.

As a data warehouse and analytics engine, either standalone or integrated with Hadoop, Exasol is being used in a wide range of Big Data use cases, including accelerating standard reporting, running multi-user ad-hoc analytics, and performing complex modelling using predictive in-database analytics.



Key capabilities

In-memory technology

Innovative in-memory algorithms enable large amounts of data to be processed in main memory for dramatically faster access times.

Column-based storage and compression

Columnar storage and compression reduces the number of I/O operations and amount of data needed for processing in main memory and accelerates performance.

Massively Parallel Processing

Exasol was developed as a parallel system based on a shared nothing architecture. Queries are distributed across all nodes in a cluster using optimised, parallel algorithms that process data locally in each node's main memory.

High user concurrency

Thousands of users can simultaneously access and analyse large amounts of data without compromising query performance.

Scalability

Linear scalability lets you to extend your system and increase performance by adding additional nodes

Tuning-free database

Intelligent algorithms continuously monitor usage and perform self-tuning, optimizing system performance and minimizing administrative work.



Key capabilities

Fast access to all data sources

Exasol's data virtualization framework ("virtual schemas") and high-performance data import framework allow users to connect to new data sources more easily and analyze them faster than before.

Various use cases

Exasol is a very flexible solution and can be integrated into various business models. Exasol can be deployed either as a software-only solution, as an appliance or in the cloud (EXACloud, Microsoft Azure or Amazon Web Services).

Comprehensive Hadoop integration

Data that is in any native format supported by HCatalog can be loaded directly from HDFS, which makes high-speed analysis of structured and unstructured data easy and straightforward. Data is transferred fast and in a parallelized manner.

Advanced In-Database Analytics

User Defined Functions (UDF) allow in-database advanced analytics to be easily run using R, Python, Lua and Java.

Advanced connectivity

In addition to existing JDBC, ODBC and .NET interfaces, Exasol now also supports a web socket-based SQL interface. Using this new interface, almost any platform can easily access Exasol, even if there is no dedicated driver available. Exasol already offers a Python adapter based on this API.



What is Exasol?

Core architecture

Exasol's architecture can be seen in Figure 1. Exasol offers a user-friendly, web-based graphical user interface (EXAOperation), its own cluster management system (EXAClusterOS) as well as its own storage management module (EXAStorage). Exasol supports ANSI standard SQL 2008 (including all analytical functions) as well as a large selection of the commonly-used Oracle SQL dialect. Its support of parts of Oracle SQL language set is of particular benefit when migrating from Oracle-based applications; there is minimal need, if any at all, to re-write or change the code.

Additional "hot standby" servers in the cluster guarantee redundancy and resilience. Should a server go down, one of the "hot standby" servers automatically takes over and the cluster can continue to operate. The defective server can then be removed from the cluster and replaced without taking the Exasol database offline, the new server then becomes the new "hot standby" server.

Logical Achitecture

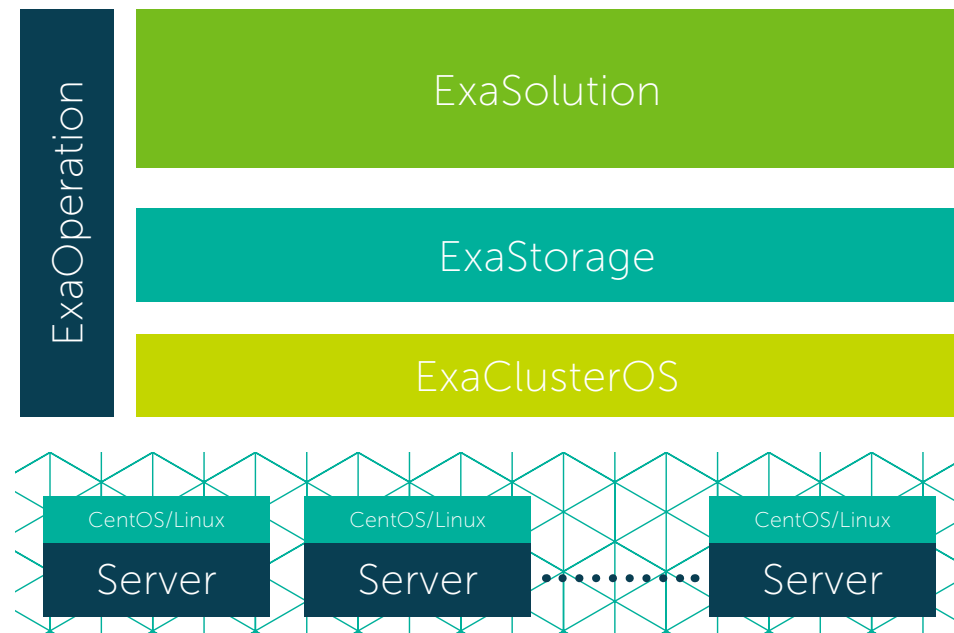


Figure 1. Exasol Logical Architecture



What is Exasol?

Core architecture

Exasol has high degrees of automation built in to its design to ensure that it delivers high performance with minimal need for expensive DBA resource to operate it. Some of the key areas of automation are:

Automatic distribution of data

evenly across all servers in the cluster

Automatic compression of data

at column-level with identical images stored both in main memory and on persistent media (hard disk) to optimize performance.

Automatic duplication of data

across servers to ensure data integrity in the event of a server failure

Automatic monitoring and logging of system resources

(RAM, Disk, CPU) to aid in capacity planning

Automatic selection of innovative compression algorithms

that are data type specific and optimized for in-memory processing. These algorithms also work independently on each node to ensure optimal performance



What is Exasol?

Core architecture

The decreasing cost of RAM has encouraged a number of vendors to develop in-memory options for their existing database products. It is important to note that Exasol has been architected and designed from the outset as an in-memory database. This is not an «add-on» feature and unlike a number of competing products, [with Exasol it is not necessary to store the entire database in memory](#).

For optimum performance, sufficient memory is required to service the query workload presented. As with persistent storage on disk, compression is also helpful in this respect. This affords considerable flexibility and advantages in terms of balancing memory, servers, performance and cost.

Since Exasol's raw performance is derived from in-memory processing, [Exasol has developed a highly intelligent, cost-based query optimizer, based on self-learning algorithms](#), that manages what data is kept in-memory. An important feature of the optimiser is the automatic creation and management of join indices in-memory,

to maximise the benefits of in-memory processing. In real-world deployments the optimiser adapts over time to the workload presented, in order to deliver optimum performance from the database resources available. Therefore, workload management at the system level is significantly automated, reducing the need for skilled and expensive DBA resource.

In addition, as part of the workload management, Exasol automatically monitors and logs its resource utilization. Therefore as workloads increase (i.e. more data, more users, more complex queries) and the performance of Exasol starts to degrade, the system monitoring information is available to help determine how much more memory to allocate to the database in each server (scale up), or if necessary, what number of new nodes (servers) need to be added to the cluster (scale out) to maintain performance levels.

Scaling up and allocating more memory to the database in each server is a straightforward command executed

through the EXAoperation GUI. Adding new servers is also a straightforward process. Once the new hardware is connected in to the cluster, data is automatically re-distributed across current and new nodes as a background process, and users can still be running queries whilst this is on-going.

In specific use-cases, where workload is highly diverse and/or user priorities vary significantly at different times of day, Exasol provides workload management through user prioritisation. This facilitates manual configuration of users' or groups of users' (roles) priorities by weighting resource allocation and, if required, defining a schedule.



What is Exasol?

Other features

Exasol supports standard interfaces for integrating upstream (Data Integration) and downstream (BI) tools. The standard interfaces that ship with the Exasol software include ODBC, JDBC, and ADO.NET. Exasol has already been tested and is working in production environments with BI Tools and Data Integration tools from all the leading vendors including: Informatica, Talend, Pentaho, Tableau, Business Objects, Cognos and Microstrategy.

Furthermore, support for the native Oracle OCI interface enables extremely fast and parallelized data exchange with Oracle database systems.

A further point of differentiation is that a SQL pre-processor allows users to transform existing proprietary queries into ANSI-compliant SQL without having to make changes to the original queries.

Exasol also has a bulk loader that is easy-to-use and compatible with data integration tools. One powerful feature lies in the ability to process compressed data, e.g. in ZIP format, that enables even faster data transfer rates. As organizations

increasingly use analytics to support their operational business, data in the data warehouse must be regularly adjusted and updated.

Exasol Advanced Edition also offers a powerful analytics framework. Users can run code that has been written in R, Python, Lua or Java as user-defined functions (UDFs) parallelized in the database with extremely high in-memory performance. In addition, any other programming languages can be integrated, so that users are no longer limited by the currently extensive standard languages that are available.

Furthermore, calculations based on the MapReduce principle in Hadoop systems can be run directly in the SQL engine and combined inside a SQL statement with standard SQL.

Exasol Advanced Edition includes a powerful data virtualization framework (virtual schemas) and an extendible and highly-flexible integration framework (ETL UDFs). This enables a very flexible and high-performance integration of Hadoop-based data in Exasol.

Another unique feature is the analytic Skyline function for preference analytics.

Preference analytics addresses the fundamental problems of traditional data mining approaches, where the everincreasing volumes of data and multiplicity of variables mean that sorting, filtering and weighting lead to sub-optimal analyses.

A real-life example is when choosing the best investment funds. Here, using conventional approaches to ensure a continuous objective analysis while taking into account risk, returns and countless other metrics is anything else but simple. Skyline calculates the usually rather small subsets of such funds, which can actually be shortlisted on grounds of the defined criteria.



Value proposition

Exasol delivers high performance analytics on a system that is easy to use, highly scalable, fast to implement and extremely cost effective.

Implementing Exasol does not require you to replace your existing systems. It can be implemented alongside existing infrastructure as a complementary solution, to provide the high performance analytics that your current systems cannot deliver. This complementary approach enables you to continue to leverage your existing investments and prove the value that high performance analytics will deliver to your business without impacting existing operations. If appropriate you can then plan a phased migration of existing analytics applications to Exasol over time.

Exasol's advantage in terms of the significant Price/Performance it delivers is clearly demonstrated by the independent Transaction Processing Performance Council TPC-H benchmark, where Exasol holds the number one position, by a significant margin over other solutions, for both raw performance and price-performance on data volumes ranging from 300GB through to 100TB.

The graph in figure 2 illustrates the performance advantage of Exasol at all scale factors and the detailed TPC-H results can be found on the Transaction Processing Performance Council website at http://www.tpc.org/tpch/results/tpch_perf_results.asp.

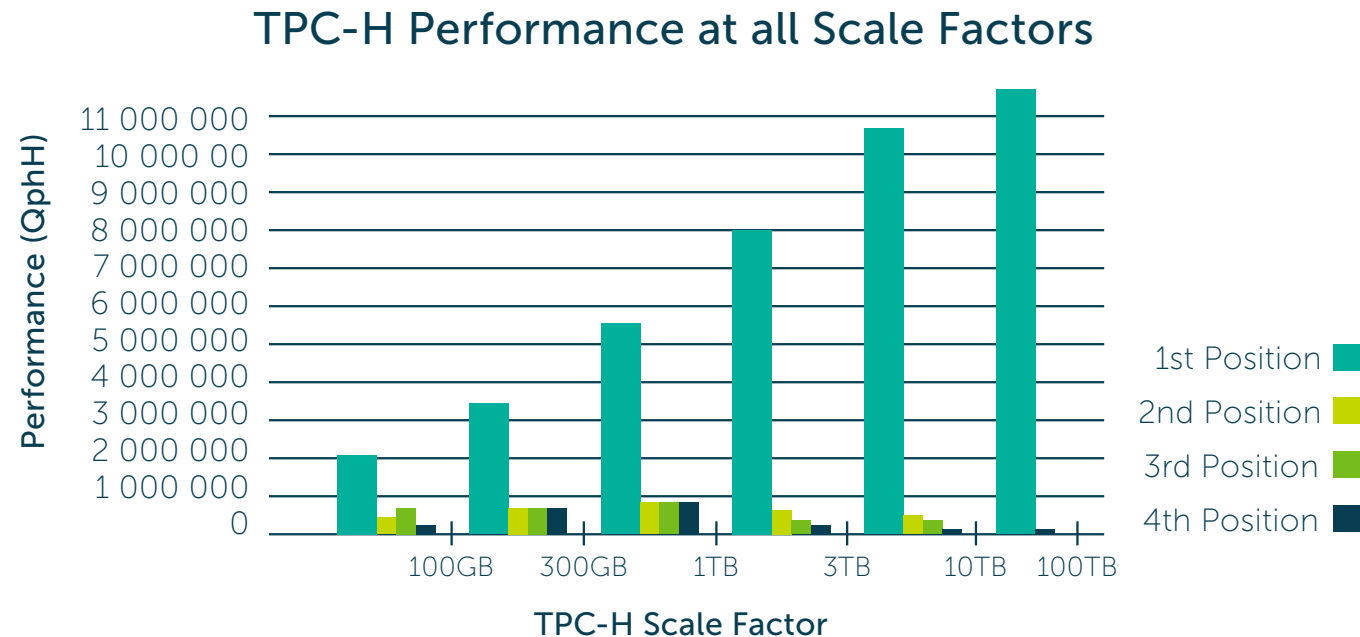


Figure 2: Exasol performance advantage

03

Value proposition

The exceptional analytics capabilities of the advanced edition open up a wide range of powerful new ways to analyze your business data. These can be roughly split into two categories:

The expansion, integration or even replacement of traditional, platform-specific analytic applications,

where the maximum amount of data that can be processed is highly limited by the platform architecture. This is often the case in systems such as MATLAB or SAS.

The creation of new high-performance computing (HPC) applications

that are only made possible with MPP technology (such as Exasol) and its scalable data analytics capabilities.

These applications are characterized by extremely large data volumes that have to be analyzed in a short time using complex algorithms.

The powerful capabilities of the advanced edition are based on open frameworks for data integration and analytical application development. The integration of these functionalities can be done easily via the standard SQL interface using UDFs (user defined functions). This open approach means that companies can now plan and build solutions that leverage existing technology investments. This differentiates Exasol from many competitors who often pursue closed, proprietary approaches that force customers into classic vendor lock-ins.



Deployment benefits

Exasol offers considerable flexibility in deployment, enabling highly efficient utilisation of both hardware and software licensing:

Exasol requires a lower specification of (commodity) hardware

to deliver a given workload;

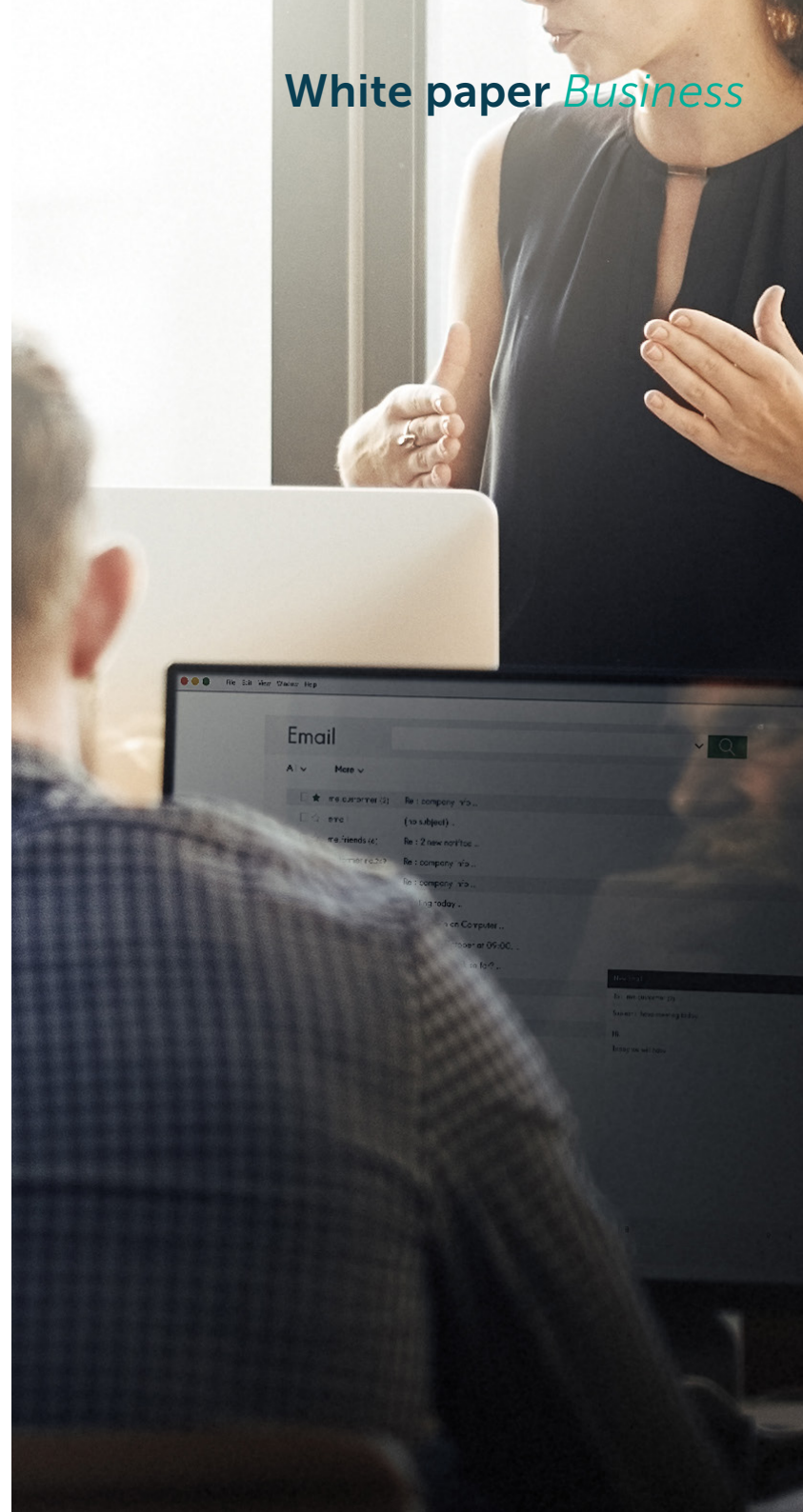
Exasol licensing is based optionally on RAM allocated to the application on the maximum amount of raw data that can be processed

– there are no end-user limits, and the storage of infrequently used or entirely new, but as yet unexplored, data sets is not penalised or the maximum amount of raw data that can be processed.

The performance of Exasol depends on the relationship between database size, system usage as well as the amount of memory available.

(Note: There is no need to ensure you have enough memory capacity for your entire data volumes.) Performance is at its most optimum when all the data needed to process queries is always in memory.

However, performance with less RAM (and lower licensing cost) typically still meets SLA requirements.



Deployment benefits

Higher workload demands can be satisfied flexibly

through either increasing memory in existing nodes (scale up) and/or adding new database nodes (scale out).

Several databases can be run on a single hardware cluster

ideal for supporting multi-tenant applications;

Exasol supports a large percentage of the Oracle language dialect,

which means that Oracle environments can be migrated quickly, efficiently and with minimal effort.

Exasol includes a SQL pre-processor

that enables applications to be migrated fast without having to customize the application itself.

Exasol's powerful analytics framework

offers a seamless integration of and with other technologies.

Exasol includes a data virtualization framework

called "virtual schemas" as well as a high performance data integration framework, so you can connect to and analyze data from more sources than ever before.



05

Conclusion

Do you want to speed up your current BI platform or business analytics? Do you want to predict complex correlations using predictive analytics? Or do you need to ensure large numbers of your users have access to BI applications and analytics in general? If you would like to offer your organization really useful “big data” solutions, then Exasol is for you – it’s the platform that allows you to deliver value fast, easily and affordably.


Take a look and see the power of Exasol for yourself – test our solution for free:

www.exasol.com/download

Or if you would like to discuss your requirements with one of our experts please contact us at

info@exasol.com





Exasol AG
Neumeyerstr. 22 – 26
90411 Nuremberg
Germany
www.exasol.com

Tel: +49 911 23991-0
Email: info@exasol.com

Follow us for the latest content:



About this Whitepaper: Information listed here may change after the data sheet has been printed (July 2018). Exasol is a registered trademark. All trademarks named are protected and the property of their respective owner. © 2018, Exasol AG | All rights reserved

