



Panel #2

LISBON
April 2026

Theme

**Is the Cloud Still Just Infrastructure?
Rethinking Data Processing in Intelligent Systems**

ComputationWorld 2026 & DataSys 2026



Speakers

LISBON
April 2026

Moderator

Prof. Dr. Christoph Neumann, OTH Amberg-Weiden, Germany

Panelists

Wolfgang Forstmeier, Siemens AG, Germany

Dr. Raju Shrestha, Oslo Metropolitan University (OsloMet), Norway

Kunal Rao, NEC Laboratories America, Inc., USA

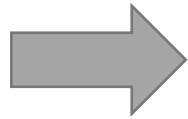
Jorge Carola, Siemens SA, Portugal



Panel Disclaimer

LISBON
April 2026

„Cloud
Databases“



“Is the Cloud Still
Just Infrastructure?
Rethinking Data Processing
in Intelligent Systems”



Christoph P. Neumann
OTH Amberg-Weiden
Germany

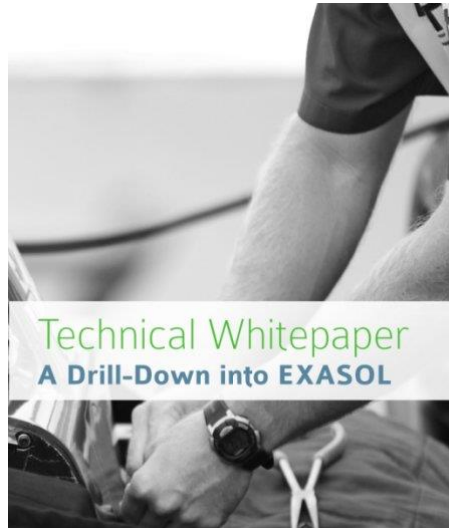


Towards Cloud Databases

(from the perspective of Analytic Databases and OLAP, i.e., for DataWarehouses, BI, and Big Data Analytics)

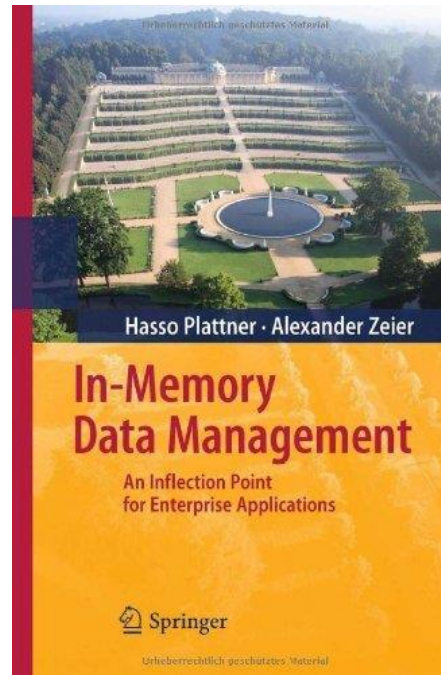
LISBON
April 2026

EXASOL 2007/2016



{Closed Source}
EXASOL

SAP / H. Plattner / HPI



{Closed Source}
SannsouciDB

{Closed Source}
HANA
(HPI „Hyrise“,
SannsouciDB?,
MaxDB, ASE/IQ?,
P*Time, Trex)

Snowflake 2016

The Snowflake Elastic Data Warehouse

Benoit Dageville, Thierry Cruanes, Marcin Zukowski, Vadim Antonov, Artin Avanes, Jon Bock, Jonathan Claybaugh, Daniel Engovатов, Martin Hentschel, Jiansheng Huang, Allison W. Lee, Ashish Motwala, Abdul Q. Munir, Steven Pelley, Peter Povinec, Greg Rahn, Spyridon Triantafyllis, Philipp Unterbrunner

Snowflake Computing

ABSTRACT

We live in the golden age of distributed computing. Public cloud platforms now offer virtually unlimited compute and storage resources on demand. At the same time, the Software-as-a-Service (SaaS) model brings enterprise-class systems to users who previously could not afford such systems due to their cost and complexity. Also, traditional data warehousing systems are struggling to fit into this new environment. For one thing, they have been designed for fixed resources and are thus unable to leverage the cloud's elasticity. For another thing, their dependence on complex ETL pipelines and physical tuning is at odds with the flexibility and business requirements of the cloud's new types of semi-structured data and rapidly evolving workloads.

We decided a fundamental redesign was in order. Our mission was to build an enterprise-ready data warehousing solution for the cloud. The result is the Snowflake Elastic Data Warehouse, or "Snowflake" for short. Snowflake is a multi-tenant, transactional, server, highly scalable and elastic system with full SQL support and built-in extensions for semi-structured and schema-less data. The system is offered as a pay-as-you-go service in the Amazon cloud. Users upload their data to the cloud and can immediately manage and query it using familiar tools and interfaces. Implementation began in late 2012 and Snowflake has been generally available since June 2015. Today, Snowflake is used in production by a growing number of small and large organizations alike. The system runs several million queries per day over multiple petabytes of data.

In this paper, we describe the design of Snowflake and its novel multi-tenant, shared-data architecture. The paper highlights some of the key features of Snowflake: extreme elasticity and availability, semi-structured and schema-less data, time travel, and end-to-end security. It concludes with lessons learned and an outlook on ongoing work.

Categories and Subject Descriptors: Data management systems; Database management systems engines

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are made or distributed for non-profit or educational advantage and for copies for the system and the full extent of the copyright notice. For other uses, contact the owner/author(s).
Snowflake Computing, Inc. 2016. All rights reserved. See <http://www.snowflake.com>.
© 2016 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-1515-5/16/04
DOI: <http://dx.doi.org/10.1145/2842961.2842971>

Keywords

Data warehousing, database as a service, multi-tenant shared data architecture

1. INTRODUCTION

The advent of the cloud marks a move away from software delivery and execution on local servers, and toward shared data centers and software-as-a-service solutions hosted by platform providers such as Amazon, Google, or Microsoft. The shared infrastructure of the cloud promises increased economies of scale, extreme scalability and availability, and a pay-as-you-go cost model that adapts to unpredictable usage demands. But these advantages can only be captured if the software itself is able to scale elastically over the pool of commodity resources that is the cloud. Traditional data warehousing solutions pre-date the cloud. They were designed to run on small, static clusters of well-behaved machines, making them a poor architectural fit.

But not only the platform has changed. Data has changed as well. It used to be the case that most of the data in a data warehouse came from sources within the organization: transactional systems, enterprise resource planning (ERP) applications, customer relationship management (CRM) applications, and the like. The structure, volume, and rate of the data were all fairly predictable and well known. But with the cloud, a significant and rapidly growing share of data comes from less controllable or external sources: application logs, web applications, mobile devices, social media, sensor data (Internet of Things). In addition to the growing volume, this data frequently arrives in schema-less, semi-structured formats [2]. Traditional data warehousing solutions are struggling with this new data. These solutions depend on deep ETL pipelines and physical tuning that fundamentally assume predictable, slow-moving, and easily categorized data from largely internal sources.

In response to these shortcomings, parts of the data warehousing community have turned to "Big Data" platforms such as Hadoop or Spark [3, 1]. While these are indispensable tools for data center-scale processing tasks, and the open source community continues to make big improvements—such as the Stripe initiative [25]—they still lack much of the efficiency and feature set of established data warehousing technology. But most importantly, they require significant engineering effort to roll out and use [26].

We believe that there is a large class of use cases and workloads which can benefit from the economics, elasticity, and service aspects of the cloud, but which are not well served by other traditional data warehousing technology or



Christoph P. Neumann
OTH Amberg-Weiden
Germany



Cloud Databases

LISBON
April 2026

Current Scientific Trends (SIGMOD, VLDB, SoCC)

- Use machine learning and large language models to augment or replace **core database components** such as query optimization, tuning, and user interaction
- Separate compute, storage, and memory into **independently scalable layers** to enable flexible, **cloud-native architectures** (Linear Scalability for NoSQL/NewSQL/MPP-DBMS traditionally by **Shared-Nothing** → Shift to **Shared-Disk 2.0**)
- Design fully managed **serverless** systems with fine-grained, **per-query elasticity** and **pay-as-you-go** execution (**auto-scaling**)
- Build self-managing databases that **automatically tune, optimize, and test** themselves with minimal human intervention (**AI Ops for DBMS**)
- Efficiently share infrastructure across **multi-tenancy** through **advanced scheduling** (scaling cloud DBs is now as much about **resource orchestration** as query processing)
- Balance high performance with **security, privacy, and compliance** in shared cloud environments (i.a., balancing with **encryption** and **isolation** in shared analytics)
- Leverage non-commodity **hardware** (e.g., GPUs, RDMA, CXL, non-volatile memory) to improve database **performance** and scalability (best paper award on CXL → Compute eXpress Link is an open standard interconnect for high-speed, high capacity CPU-to-device and CPU-to-memory connections...)
- AI-native databases (**Vector Databases** as well as **LLMs inside RDBMS**)



Christoph P. Neumann
OTH Amberg-Weiden
Germany



Venture Capitalists Realism

LISBON
April 2026

„You're crazy
if you don't start in the cloud;
you're crazy
if you stay on it.“



Christoph P. Neumann
OTH Amberg-Weiden
Germany

Sarah Wang and Martin Casado @ Andreessen Horowitz (2021)

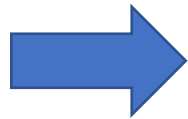
[The Cost of Cloud, A Trillion Dollar Paradox](#)



Towards Panelists

LISBON
April 2026

„Cloud
Databases“



**“Is the Cloud Still
Just Infrastructure?
Rethinking Data Processing
in Intelligent Systems”**



Christoph P. Neumann
OTH Amberg-Weiden
Germany



Towards Panelists

LISBON
April 2026

Is the Cloud Still Just Infrastructure? Rethinking Data Processing in Intelligent Systems



Radio Edit
(~ 2½ minutes)



Extended Mix (8 minutes) with Lyrics

<https://suno.com/s/EPWfpZ2DxXLXybwM>



Christoph P. Neumann
OTH Amberg-Weiden
Germany



Panelist Position

LISBON
April 2026

Cloud as centralized infrastructure element

- **Data is massively distributed**
 - Data in industrial context is usually extremely distributed with doubtful internet connection
 - Sending all data to the cloud is economically nonsense
 - Data processing results are frequently required where the data originates
- **The world trends to de-globalization**
 - A lot of different data privacy rules apply nowadays
 - Not all cloud providers can be used anywhere
 - Data and digital sovereignty is becoming a more popular topic



Wolfgang Forstmeier
Siemens AG
Germany



Panelist Position

LISBON
April 2026

Cloud as centralized infrastructure element

- **AI and Agentic workflows**
 - Re-evaluate the Software as a Service (SaaS) model when the primary "users" are no longer humans but intelligent agents or AI
 - Inferencing tends to be a more distributed mechanism
 - Training of models needs heavy compute and large datasets
 - More event driven and coordination tasks for cloud environment
- **Combining Virtual and Physical world**
 - Inferencing needs real-time capabilities
 - Local execution of models due to autonomy and resilience requirements
 - Data is only valid for a very short period (movement of robot arms)
 - Contextualization, re-training, training, can/should be done in a higher-level based system.



Wolfgang Forstmeier
Siemens AG
Germany



Panelist Position

LISBON
April 2026

The cloud is no longer just infrastructure!

It is becoming the operating layer of intelligent systems



Raju Shrestha
OsloMet – Oslo Metropolitan
University, Norway

➤ **It orchestrates intelligences**

- Not just compute and storage anymore
- Coordinates data, models, and decisions

➤ **Distributed processing is the new normal**

- Intelligence runs across device, edge, and cloud
- Placement depends on latency, cost, and privacy

➤ **Data pipelines are now intelligence pipelines**

- From ETL (extract, transform load) to real-time stream processing
- Continuous ingestion, inference, and feedback

➤ **AI makes the cloud active**

- Optimizes scaling, routing, and scheduling
- The cloud help shape system behavior



Panelist Position

LISBON
April 2026

We must rethink cloud architecture for intelligent systems

➤ Architecture must support the full loop

- Sense, process, infer, learn, respond

➤ Governance must be built in

- Data lineage, sovereignty, trust, explainability

➤ Edge + cloud is the winning model

- Local responsiveness + global coordination

➤ Metrics must expand

- Performance, cost, energy, privacy, resilience; not uptime alone

The cloud is no longer merely infrastructure - it is the adaptive control layer where data, intelligence, and governance converge.



Raju Shrestha
OsloMet – Oslo Metropolitan
University, Norway



Panelist Position

LISBON
April 2026

- **Cloud today is more than just infrastructure !**
 - It provides all necessary building blocks to build AI-native and traditional/non-AI applications
 - It serves as an end-to-end orchestration platform with support for auto-scaling
 - Latest trends and developments in application engineering (e.g. new LLM models, RAG, etc.) are immediately made available to use
 - It keeps improving behind the scene
 - However, most, but not all need it. On-prem still remains relevant for regulated and latency-critical systems



Kunal Rao
NEC Laboratories America, Inc.
Princeton, NJ, USA



Panelist Position

LISBON
April 2026

Cloud as Intelligence Platform

- **Distributed Processing**
 - Moving from centralized processing to edge and hybrid architectures
 - Field devices are also part of the picture now and can be cloud orchestrated and coordinated with other resources and not just fleet-managed
 - Modern Edge/Field devices have increased processing power, allowing local processing to avoid the costs of moving raw data, especially in large fleets
- **Trust and reliability**
 - Edge-to-cloud continuum increase the trust in the systems and improves the reliability
 - Synchronized workload migration with zero downtime on a edge-to-cloud continuum increases the system reliability specially in catastrophic scenarios
- **Coordination of intelligent systems**
 - Cloud has system optimizer
 - Cloud has central coordinator for intelligent systems, not only for training but acting as task intelligent store



Jorge Carola
Siemens SA
Portugal