CUMULONIMBO BigData Platform. Analytics in a wink

The Power of a Wink: while you wink we answer your queries over terabytes of data, process millions of transactions and correlate millions of events

Contact information

Address: ETSI Informáticos – UPM, Campus de Montegancedo, 28660 Boadilla del Monte (Madrid)
Phone number: 910672801
Website: fi.upm.es
Email: mpatino@fi.upm.es

Technological Offers type

Technological solutions

Research and innovation areas

- Digital Technologies, Artificial Intelligence, Cybersecurity, 5G, Robotics

ODS

Available from: 2020

Where?

Artificial Intelligence R&I Centre (AI.nnovation Space) Distributed Systems Laboratory (DSL)

Keywords: |

Brief description of the technology solution and the added value it provides

An ultra-scalable transactional database with On Line Analytical Processing (OLAP) capabilities able to query TBs of data in seconds while processing million transactions per second and correlating millions of events per second. CumuloNimbo BigData platform is a scale-out solution for OLTP, OLAP and CEP where you store your data once therefore, avoiding costly ETLs and its associated headaches. It also provides transparent scalability for transactional processing on commodity hardware enabling to scale the database in public or private clouds avoiding sharding. The solution is built on top of NoSQL technology, Hadoop HDFS and HBase providing also full ACID transactions for HBase and enabling to combine SQL and NoSQL as needed. The integration with the Hadoop ecosystem enables to exploit existing open source products from this ecosystem such as R for predictive analysis and Mahout for large scale machine learning.

Description of the technological base

It is the first solution able to scale transactions providing full ACID transactions with full SQL support and providing standard interfaces (JDBC, ODBC) in a fully transparent way to applications (your data management application will be scaled without touching a single comma) based on commodity hardware and amenable to be run in public and private clouds.

The platform also incorporates a scale-out parallel query processing to provide OLAP support and perform analytical queries over TBs of data in seconds.

Finally, the platform also integrates a scale-out parallel complex event processing engine able to correlate millions of events per second.

In a wink, we answer your queries over TBs of data, process millions of transactions updating your data, and correlate millions of events with your data

Market demands

- **Scalable Databases in Public & Private Clouds:** Current cloud databases do not scale. They way to overcome this gap is by means of sharding, splitting the database into fragments. However, the transactional ACID properties are lost extremely complicating the application requiring very costly changes.
- **Avoiding copying data through ETLs:** Current database systems are specialized on different tasks. Transactional databases are used for the production database due to their ability to support updates in a consistent manner. Data warehouses are used for getting online responses to heavy analytical queries. However, this approach requires to architect, develop and plan a process to copy from the production database to the data warehouse, the so-called ETL process. It is estimated that ETLs are 75-80% of business analytics.
- **Operational Intelligence and New Applications:** Many companies have the same recurrent question: how to monitor their business processes to react faster to opportunities and threats. Current IT infrastructure results in reaction times of days.

Competitive advantages

- Scalable transactions for public & private clouds providing full ACID transactions & full SQL transparently to applications: Avoids sharding and 100% of its cost.
- OLTP+OLAP on a single database: Avoids the costs of ETLs that are estimated to be 75-80% of the overall costs of analytics. Enables real-time business intelligence.
- Correlate massive events with your data: Enabling operational intelligence with OLTP+CEP+OLAP
- Based on commodity hardware and scaling out: Avoids the costs of appliances, usable in clouds.
- Based on Hadoop ecosystem and interoperable with ecosystem tools such as R, Mahout, Flume, etc.

"Scale your Database Without Limits Avoiding Sharding and its Associated High Cost"

Previous references

- A large international bank is already evaluating the adoption of the platform.
- The co-founders have led 5 European projects and obtained and managed 6+ Million Euro in the last 5 years. They have been intensively researching scalable data management during the last 15 years.

Intellectual property

2 filed patents, protecting all the core innovations:

- USPTO 61/356,353 (scalable transactions)
- USPTO 61/561/508 (parallel query processing)

Development stage

- Concept
- R&D
- Lab-Prototype
- Industrial Prototype
- **Production**

Contact

Contacto CumuloNimbo BigData Platform

Ricardo Jiménez-Peris, Marta Patiño-Martínez

Distributed Systems Lab (LSD) – UPM

e: {rjimenez, mpatino}@fi.upm.es

Contacto UPM

Área de Innovación, Comercialización y Creación de Empresas

Centro de Apoyo a la Innovación Tecnológica – UPM

e: innovacion.tecnologica@upm.eS