

CUMULONIMBO BIGDATA PLATFORM

Análisis de datos en un parpadeo.

El poder de un parpadeo: mientras parpadeas, respondemos a consultas sobre terabytes de datos, procesamos millones de transacciones y correlacionamos millones de eventos.



Información de contacto

Dirección: ETSI Informáticos - UPM, Campus de Montegancedo, 28660 Boadilla del Monte (Madrid)

Teléfono: 910672801

Página web: fi.upm.es

Correo electrónico: mpatino@fi.upm.es

- [Consultar disponibilidad](#)

Tipo de oferta tecnológica

Soluciones tecnológicas

Áreas de investigación e innovación

- Tecnologías digitales, Inteligencia Artificial, ciberseguridad, 5G, robótica

ODS



Disponible desde: 2020

¿Dónde?

Centro de I+D+i en Inteligencia Artificial (AI.nnovation Space) Laboratorio de sistemas distribuidos (LSD)

Palabras clave: | [bases de datos](#)

Descripción breve conjunta de la solución y valor añadido que aporta

Una base de datos transaccional ultra-escalable con capacidades OLAP (On Line Analytical Processing) capaz de realizar consultas de TBs de datos en segundos mientras procesa millones de transacciones por segundo y correlaciona millones de eventos por segundo. La plataforma CumuloNimbo BigData es una solución escalable para OLTP, OLAP y CEP en la que se almacenan los datos una sola vez, evitando, por lo tanto, costosos ETLs (Extract, Transform and Load) y los dolores de cabeza asociados. También proporciona una escalabilidad transparente para procesamiento transaccional en hardware que permite escalar la base de datos en nubes públicas o privadas, evitando sharding. La solución se construye sobre tecnología NoSQL, Hadoop HDFS y HBase proporcionando también transacciones ACID para HBase y permitiendo combinar SQL y NoSQL según las necesidades. La integración con el ecosistema de Hadoop permite explotar productos open-source como R para análisis predictivo y Mahout para técnicas machine learning a gran escala.

Descripción de la base tecnológica

Es la primera solución para escalar transacciones permitiendo transacciones ACID con soporte total para SQL y proporcionando interfaces estándar (JDBC, ODBC) de forma totalmente transparente para aplicaciones (tu aplicación de gestión de datos escalará sin ningún tipo de intervención por tu parte) basadas en hardware no especializado y susceptible de ser ejecutados en clouds públicas y privadas.

La plataforma también incorpora un procesamiento paralelo de consultas escalable para proporcionar soporte OLAP y realizar consultas analíticas sobre TBs de datos en cuestión de segundos.

Por último, la plataforma también integra un motor paralelo escalable de procesamiento de eventos complejos capaz de correlacionar millones de eventos por segundo.

"En lo que dura un parpadeo, respondemos consultas sobre TBs de datos, procesamos millones de transacciones actualizando sus datos, y correlacionamos millones de eventos con sus datos"

Necesidades de negocio / aplicación

Bases de datos escalables en Clouds públicas y privadas

Las actuales bases de datos en cloud no escalan. La forma de superar esto es a través de sharding, dividiendo la base de datos en fragmentos. Sin embargo, de esta forma, las propiedades transaccionales ACID se pierden, complicando el mantenimiento de las aplicaciones y requiriendo cambios muy costosos.

Evitar copiar datos a través de ETLs

Los sistemas actuales de bases de datos están especializados en diferentes tareas. Las bases de datos transaccionales se usan para la creación de bases de datos debido a su capacidad para permitir actualizaciones de forma consistente. Los repositorios de datos se usan para obtener respuestas online a consultas analíticas de gran tamaño. Sin embargo, este enfoque precisa de diseñar, desarrollar y planificar un proceso para copiar de las bases de datos de producción a los repositorios de datos (los llamados ETLs). Se estima que los ETLs son el 75-80% del business analytics.

Inteligencia operacional y nuevas aplicaciones

Muchas empresas se plantean una pregunta recurrente: cómo monitorizar sus procesos de negocio para reaccionar más rápidamente a oportunidades y amenazas. Las actuales infraestructuras TI terminan por reaccionar en órdenes de días.

Ventajas competitivas

- Transacciones escalables para clouds públicos y privados proporcionando transacciones ACID al 100% y soporte SQL de forma transparente para las aplicaciones: evita sharding y el 100% de su coste.
- OLTP+OLAP en una única base de datos: evita los costes de ETLs que se estiman entre el 75-80% de los costes globales de las prácticas de analytics. Permite inteligencia de negocio en tiempo real.
- Correlaciona eventos a escala masiva con tus datos: permite inteligencia operacional con OLTP+CEP+OLAP.
- Basado en hardware no específico y escalabilidad: evita los costes de appliances, usable en clouds.

- Basado en el ecosistema Hadoop e interoperable con herramientas de dicho ecosistema, tales como R, Mahout, Flume...

“Escala tu base de datos sin límites evitando sharding y su alto coste asociado”

“Realiza consultas analíticas en tu base de datos transaccional en producción evitando costosos ETLs”

“Correlaciona millones de eventos por segundo”

Referencias

Los co-fundadores han liderado 5 proyectos europeos y gestionados por ellos más de 6 M euros en los últimos 5 años. Han investigado de forma intensiva en gestión escalable de datos en los últimos 15 años.

Protección industrial

- Patente concedida en EEUU US20160179876.
- Patente europea concedida vía EPO EP2780832.
- Patente concedida en España ES2584102.

Grado de desarrollo

- Concepto
- Investigación
- Prototipo-Lab
- Prototipo industrial
- **Producción**

Contacto

Contacto CumuloNimbo BigData Platform

Ricardo Jiménez-Peris, Marta Patiño-Martínez

Distributed Systems Lab (LSD) – UPM

e: mpatino@fi.upm.es

Contacto UPM

Programas de Innovación y Emprendimiento

Centro de Apoyo a la Innovación Tecnológica – UPM

e: innovacion.tecnologica@upm.es