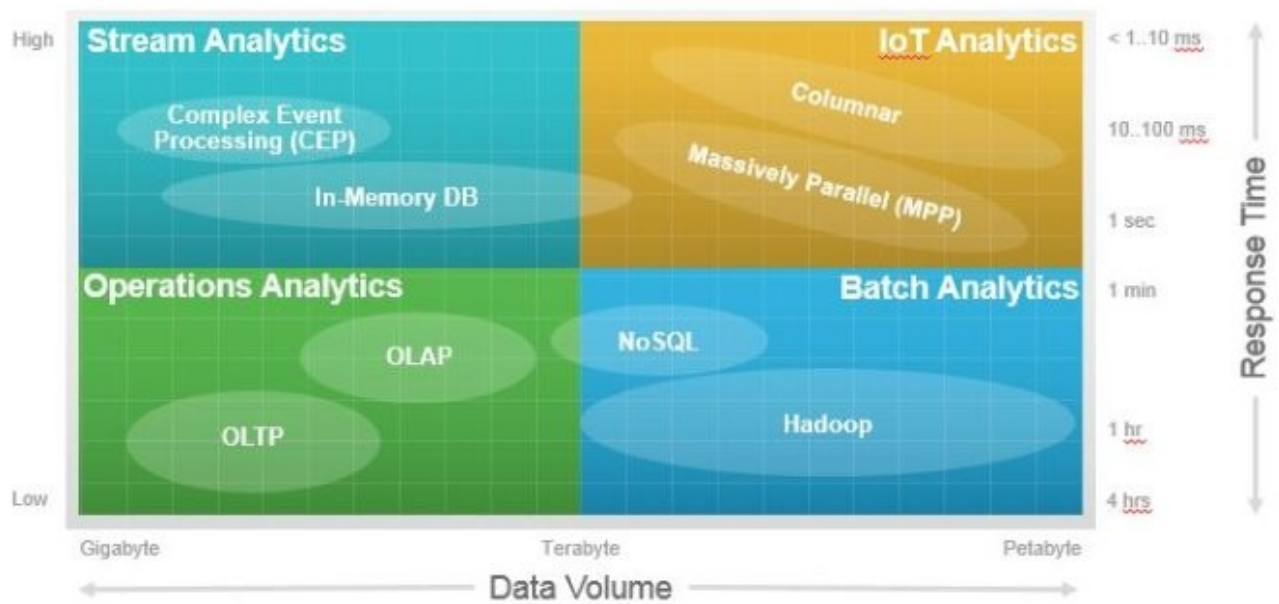




Choose Your Database Based on Your Analytics



What database do you choose for analytics?

...

Shankar Meganatha | [Follow](#)
Enterprise Cloud Sales Architect

25 2 1

Megan

Choosing the right database for analytics is like picking the right life partner. If you have chosen the right one, you are likely going to find your life happy, fulfilling, and rewarding. On the contrary, if you have chosen the wrong one, you are going to find your life miserable, stuck, ruined. Choosing the right database for your analytics is no different and should be no accident.

An analytics database is specifically designed to support business intelligence and analytics applications and to address the needs of analyzing large volumes of data quickly and faster than transactional databases. Transactional databases that are meant to support and manage your day-to-day business operations don't lend themselves for analytics.

There are different types of databases for different types of analytics. The choice of database for your analytics depends on the following criteria: 1) the type of data being analyzed, 2) the amount of data that needs to be analyzed, 3) how fast the data needs to be analyzed, and 4) IT resources and team's skill sets. The data that needs to be analyzed can be either structured or unstructured, the amount of data that needs to be analyzed can be from terabytes to petabytes, the data that needs to be analyzed in real-time for immediate insights or can have



Operations Analytics Databases

Operations Analytics is the type of analytics that helps companies maximize their business productivity and profitability. Relational and multi-dimensional databases are the most common databases for Operations Analytics. Relational databases store data in rows and columns and they include Microsoft SQL Server, Oracle, Sybase, DB2, Informix, MySQL, etc. Relational databases can be either OLTP (On-line Transactional Processing) for transactional implementations or OLAP (On-line Analytical Processing) for analytical implementations. Multi-dimensional databases store and manage data in dimensional arrays, indexed by dimensions and measured over time. Some examples of multi-dimensional databases include Oracle Essbase, SAS, IBM Cognos, Microsoft Analysis Services, etc. Since multi-dimensional database capabilities are now included in traditional relational databases and/or replaced with newer technologies, they are not currently as relevant as they used to be 10-15 years ago.

Batch Analytics Databases

Batch Analytics is the type of analytics used to process high volumes of data where a group of transactions (such as social media, web log data) is collected over a period of time. They are not essential for fast analytics necessary for real-time decision making. Big Data technologies such as Hadoop, NoSQL (Not Only SQL) are most suitable for Batch Analytics databases. Hadoop is suitable for processing of large volumes of colder data for analytics purposes. To process large volumes of data, Hadoop manages the distribution of processing in parallel, and typically across many servers by splitting the dataset into smaller sets that is called MapReduce. Top Hadoop vendors include Cloudera, MapR, HortonWorks, IBM, Amazon. NoSQL databases can be either key-value stores, wide-column stores, or document databases. NoSQL database is suitable for operational big data loads such as mobile, web applications. Examples of NoSQL database include Cassandra, MongoDB.

Stream Analytics Databases

Stream Analytics is the analysis of high velocity in-motion data and event stream that can occur as the result of set of actions such as financial transaction, equipment failure, or some other event and condition based triggers. In-memory database is one of the common databases used for stream analytics. An in-memory database primarily relies on main memory (contrary to relational databases where data is stored in disks) for storing data, processing, computing to facilitate fast response times. SAP HANA, VoltDB, TimesTen are some of the in-memory database systems available for stream analytics. The three leading incumbent relational database vendors Oracle, IBM and Microsoft also offer some in-memory capabilities in the latest versions of their databases that companies can take advantage without migrating the applications to another database.

IoT Analytics Databases

IoT (Internet of Things) Analytics refers to processing and analyzing data collected from sensor, devices, and gateways at the edge. This type of analytics is applied to industrial automation, manufacturing, and oil and gas equipment for predictive maintenance and real-time asset optimization. Columnar and Massively Parallel Processing (MPP) databases are



S'identifier

S'inscrire

Columnar databases are very efficient for queries that involve only a few columns and can aggregate queries against vast amount of data. They employ sophisticated compression and unique indexing capabilities allowing them to retrieve and analyze millions of records in billions of records with sub-second response times. [Cisco ParStream](#), Sybase IQ, Vertica are some examples of columnar databases. Massively Parallel Processing (MPP) database is a type of architecture where data is partitioned across multiple servers or nodes with each one having its memory/processors to process data locally (shared nothing). Teradata, Netezza are two leading MPP technologies and they can be used for other types of analytics besides IoT Analytics.

The database choices noted above for analytics only briefly touch upon some general aspects. Some databases can be used for more than one type of analytics and one analytics can be performed by two or more databases. It also comes down to what engineering resources exist within your company. With some engineering resources you have more choice, and with others you have only one choice. If you find yourself that you are using the wrong database for the wrong analytics, it is worthwhile to consider migrating to the right database to improve the efficiency (both cost and operational), performance and scalability of analytics applications despite possible challenges and risks that may come with any database migration.



Shankar Meganatha

Shankar Meganatha
Enterprise Cloud Sales Architect

Follow

2 comments



article-

comment__guest-

image

[Sign in](#) to leave your comment

Ильдар Гараев

Ильдар Гараев DBA Oracle – KDB bank Uzbekistan

10 mois

Гараев

бесплатный ClickHouse от Яндекса,

J'aime Répondre



Suraj Kumar Talreja

Data Analyst at Standard Chartered Bank | Banking | Insurance | Tableau Certified

2 ans

Talreja

Thanks Sir for writing such a nice article. Kudos !!

J'aime Répondre

[S'identifier](#)[S'inscrire](#)

The Essence of IoT System

The Essence of IoT System

April 12, 2017

Analyzing the Analytics types in an Ever-increasing Analytical World

Analyzing the Analytics Types in an...

February 1, 2017

What platform needed for business marketplace?

What platform are needed

December 2



© 2020

À propos

[Accessibilité](#)

[Conditions générales d'utilisation de LinkedIn](#)

[Politique de confidentialité](#)

[Politique relative aux cookies](#)

[Politique de copyright](#)

[Politique de la marque](#)

[Réglages invités](#)

[Directives de la communauté](#)

[Langue](#)