Database Management Systems

|  |  |
| --- | --- |
| 1. DBMS name | MySQL |
| Owner | Oracle Corporation |
| Supported Models | Relational |
| Who is using it (min 3) | Google, Facebook, Adobe |
| Availability tools and how it works | Replication in multiple nodes through a two-phase commit mechanism. |
| Data partitioning and how it works | Yes, auto-sharding based on a hashing algorithm. It supports cross-shards queries and transactions. |
| On-Premise, on-cloud or hybrid | On-premise, On-cloud |
| Data manipulation language | SQL |
| Data Storage System | File spaces divided into pages (InnoDB) |
| Other Interesting Features | It is the most secure and reliable DBMS. It has on-demand scalability. |

|  |  |
| --- | --- |
| 2. DBMS name | PostgreSQL |
| Owner | PostgreSQL Global Development Group |
| Supported Models | Object-relational |
| Who is using it (min 3) | OpenStreetMap, Instagram, TripAdvisor |
| Availability tools and how it works | Hot Standby/Streaming Replication provides asynchronous replication to one or more standbys. They can be queried as read-only databases. |
| Data partitioning and how it works | Yes, partitioning through table inheritance where partitions are child tables. Supports range and list partitioning. |
| On-Premise, on-cloud or hybrid | On-premise, On-cloud |
| Data manipulation language | SQL |
| Data Storage System | Array of pages, heap files |
| Other Interesting Features | No licensing cost, extensible, standards compliance. |

|  |  |
| --- | --- |
| 3. DBMS name | SQLite |
| Owner | D. Richard Hipp |
| Supported Models | Relational |
| Who is using it (min 3) | Adobe Systems, Evernote, Ruby on Rails |
| Availability tools and how it works | No native replication |
| Data partitioning and how it works | No |
| On-Premise, on-cloud or hybrid | On-premise |
| Data manipulation language | SQL |
| Data Storage System | Database file with multiple pages, b-trees |
| Other Interesting Features | It isn’t a client/server engine. Emphasis on efficiency, economy, independence and simplicity. Suitable for the internet of things. |

|  |  |
| --- | --- |
| 4. DBMS name | SQL Server |
| Owner | Microsoft |
| Supported Models | Relational |
| Who is using it (min 3) | Dell, Yahoo, Microsoft |
| Availability tools and how it works | Transactional replication (server to server), merge replication (mobile and distributed applications) and snapshot replication (for initial data sets). |
| Data partitioning and how it works | Table and index partitioning (horizontal partition, treated as a single logical entity). |
| On-Premise, on-cloud or hybrid | Hybrid |
| Data manipulation language | SQL |
| Data Storage System | Data files with pages |
| Other Interesting Features | Many additional tools such as SQL Server Management Studio and BI Tools. Ease of use and well documented. |

|  |  |
| --- | --- |
| 5. DBMS name | Oracle Database |
| Owner | Oracle |
| Supported Models | Relational |
| Who is using it (min 3) | Coca-Cola Femsa, Capcom, Chevron-Texaco |
| Availability tools and how it works | Replication through Oracle Streams. All changes are captured and pushed to a queue in each database, then the changes are applied. |
| Data partitioning and how it works | Different partitioning methods are available, using a hash, range or list to determine how to split the rows. |
| On-Premise, on-cloud or hybrid | On-Premise, On-cloud, Hybrid |
| Data manipulation language | SQL |
| Data Storage System | Data files, control files and online redo logs |
| Other Interesting Features | Support for more than 20 programming languages, stability and reliability. |

|  |  |
| --- | --- |
| 6. DBMS name | DB2 |
| Owner | IBM |
| Supported Models | Relational, object-relational |
| Who is using it (min 3) | Vantiv, RocketOn, Fifth Third Bank |
| Availability tools and how it works | Replication is available through the Replication Center GUI or through automated scripts. |
| Data partitioning and how it works | Allows database partitioning, table partitioning and organizing rows (multidimensional clustering). |
| On-Premise, on-cloud or hybrid | On-premise, On-cloud |
| Data manipulation language | SQL |
| Data Storage System | Storage groups with a set of volumes, table spaces divided into pages. |
| Other Interesting Features | It has a powerful dialect of SQL. It has a feature to automatically change memory allocation depending on the workload. |

|  |  |
| --- | --- |
| 7. DBMS name | SAP HANA |
| Owner | SAP SE |
| Supported Models | Column-oriented relational |
| Who is using it (min 3) | Cisco, HP, Lenovo |
| Availability tools and how it works | Synchronous and asynchronous replication. The synchronous mode waits for confirmation from the secondary system to consider a log write as successful. |
| Data partitioning and how it works | Allows database and table partitioning (vertical). Optimizes the use of the available memory on each node. |
| On-Premise, on-cloud or hybrid | On-premise, On-cloud |
| Data manipulation language | SQL |
| Data Storage System | In-memory |
| Other Interesting Features | Performs advances analytics and has ETL capabilities. Suitable for Big Data. |

|  |  |
| --- | --- |
| 8. DBMS name | Redis |
| Owner | Salvatore Sanfilippo |
| Supported Models | Key-value |
| Who is using it (min 3) | Twitter, Github, Snapchat |
| Availability tools and how it works | Master-slave replication. The master sends a stream of commands to keep the slave up to date. |
| Data partitioning and how it works | Yes, Redis Cluster supports range and hash partitioning. |
| On-Premise, on-cloud or hybrid | On-premise, On-cloud |
| Data manipulation language | Redis commands |
| Data Storage System | In-memory, RDB and AOF files for persistence |
| Other Interesting Features | Persistent on disk. Values can be complex data types. |

|  |  |
| --- | --- |
| 9. DBMS name | Riak |
| Owner | Basho Technologies |
| Supported Models | Key-value |
| Who is using it (min 3) | Bet 365, Best Buy, The Weather Channel |
| Availability tools and how it works | Intelligent replication. It uses a masterless, multi-node architecture. |
| Data partitioning and how it works | Each node in the cluster has a set of virtual nodes that store a different portion of the keys. |
| On-Premise, on-cloud or hybrid | On-Premise, On-cloud, Hybrid |
| Data manipulation language | Solr full-text search, secondary indexes and Map Reduce |
| Data Storage System | In-memory (log-structured hash tables), on-disk (SSTable) |
| Other Interesting Features | Global Object Expiration to remove aged data from the database. Real-time analytics with Apache Spark. |

|  |  |
| --- | --- |
| 10. DBMS name | Apache Cassandra |
| Owner | Apache Software Foundation |
| Supported Models | Column-based |
| Who is using it (min 3) | Apple, Netflix, Reddit |
| Availability tools and how it works | Automatic replication to multiple nodes for fault-tolerance. Replication across data centers is also supported. |
| Data partitioning and how it works | Two partitioners: order preserving (keeps similar keys close to each other) and random (achieves a better load balancing). |
| On-Premise, on-cloud or hybrid | On-premise, On-cloud |
| Data manipulation language | Cassandra Query Language |
| Data Storage System | Memtables and SSTables |
| Other Interesting Features | Decentralized model, every node can service a given request. Highly scalable. |

|  |  |
| --- | --- |
| 11. DBMS name | CouchDB |
| Owner | Apache Software Foundation |
| Supported Models | Document-based |
| Who is using it (min 3) | Credit Suisse, BBC, Canal+ |
| Availability tools and how it works | Bi-directional replication. Multiple nodes can have their own copies of the same data and then changes are synchronized. |
| Data partitioning and how it works | Hash partitioning through CouchDB Lounge. Every node is allocated a portion of the hash, all nodes get an equal load. |
| On-Premise, on-cloud or hybrid | On-premise, On-cloud |
| Data manipulation language | Javascript |
| Data Storage System | JSON |
| Other Interesting Features | Multiversion concurrency control (the database is not locked during writes). Devices can go offline and sync when they’re back online. |

|  |  |
| --- | --- |
| 12. DBMS name | MongoDB |
| Owner | MongoDB Inc. |
| Supported Models | Document-based |
| Who is using it (min 3) | Ebay, MetLife, Expedia |
| Availability tools and how it works | Replica sets with two or more copies of the data. All writes and reads are done on the primary replica and when it fails, a secondary replica becomes the primary. |
| Data partitioning and how it works | Sharding where each shard contains a subset of the data and a query router called *mongos* provides an interface between client applications and the sharded cluster. |
| On-Premise, on-cloud or hybrid | On-premise, On-cloud |
| Data manipulation language | Document-based query language |
| Data Storage System | BSON |
| Other Interesting Features | It can be used as a file system to store files. User-defines Javascript functions can be used to query the database. |

|  |  |
| --- | --- |
| 13. DBMS name | Neo4J |
| Owner | Neo Technology |
| Supported Models | Graph-based |
| Who is using it (min 3) | Walmart, Airbnb, NASA |
| Availability tools and how it works | The High Availability Cluster is a master-slave replication tool. Data created on the master is pushed to the slaves. The number of slaves and the frequency of the updates can be configured. |
| Data partitioning and how it works | No graph partitioning |
| On-Premise, on-cloud or hybrid | On-premise, On-cloud |
| Data manipulation language | Cypher Query Language |
| Data Storage System | Linked lists of fixed-size records |
| Other Interesting Features | Nodes and edges can have attributes and labels. Performance stays constant as data grows. |

|  |  |
| --- | --- |
| 14. DBMS name | Couchbase |
| Owner | Couchbase Inc. |
| Supported Models | Document-based, key-value |
| Who is using it (min 3) | BD, Nielsen, Betfair |
| Availability tools and how it works | Couchbase distinguishes between active data (written by a client) and replica data (copies). After a write, the data is copied to another node. Then, the replica data is distributed in the cluster to prevent a single point of failure. |
| Data partitioning and how it works | Partitioning is done by separating the data into several buckets (data containers) that can be in RAM or on disk. |
| On-Premise, on-cloud or hybrid | On-premise, On-cloud |
| Data manipulation language | N1QL |
| Data Storage System | JSON |
| Other Interesting Features | Memcache buckets cache frequently used data in RAM to reduce the number of queries. |

|  |  |
| --- | --- |
| 15. DBMS name | ArangoDB |
| Owner | ArangoDB GmbH |
| Supported Models | Document-based, key-value, graph-based |
| Who is using it (min 3) | Thomson Reuters, Liaison, AboutYou |
| Availability tools and how it works | Provides synchronous and asynchronous replication. The first one waits for all replicas to write the data before approving the write operation. The second one uses a write-ahead log. |
| Data partitioning and how it works | Sharding is supported across several primary DBServers (where the data is hosted). A hash is used to determine in which shard the data is stored. |
| On-Premise, on-cloud or hybrid | On-premise, On-cloud |
| Data manipulation language | ArangoDB Query Language, GraphQL |
| Data Storage System | JSON |
| Other Interesting Features | It has three NoSQL models in a single database. It works in a distributed cluster, unlike other graph-based databases. |

References

Digital Ocean. “A Comparison of NoSQL Database Management Systems and Models”. *Digitalocean.com.* February 21, 2014. Web. January 16, 2018.

Mullins, Craig S. “Which relational DBMS is best for your company?”. *Techtarget.com*. April 2015. Web. January 16, 2018.

MySQL. “Why MySQL?”. *Mysql.com*. 2018. Web. January 10, 2018.