Remy Lagrois

Quiz 13

Section 204

1. **b) Not only SQL** Not only SQL means that there are ways other than SQL to use and structure a database that more closely match the reality of modern data storage.
2. **c) A distributed database can optimize only two of C, A or P simultaneously.** You can try and make the third as good as possible but not perfect. Trying to maintain consistency may limit partition tolerance while maintaining partition tolerance will reduce availability. Maintaining availability will reduce partition tolerance.
3. **a) Managing consistency** A database must keep data consistent across the whole database. If information is changed in one place it must also be changed in the other places the same information is stored. In NoSQL it’s more important that consistency be made eventually instead of instantaneously.
4. **d) Availability** The information should be accessible whenever it’s needed. This means that both the hardware and software running a database must be reliable.
5. **c) Partition toleration** refers to the ability for a system to continue to run and process data in the event of a node failure or loss of communication between nodes.
6. **a) MongoDB** uses JSON like documents with flexible schemas.
7. **d) Neo4j** is the most popular graph database.
8. **b) Cassandra** is a popular column family database which distributes large amounts of data over many servers allowing for reliability and high scalability.
9. **c) Riak** Riak is a key-value database operating under a freemium model and uses the principle of eventually consistent.
10. **a) the relationships and their properties** Graph databases are very good at showing the connections and relationships between what is being stored in the database using nodes and edges.
11. **b) Improved ability to keep data consistent.** Consistency is less important in a NoSQL database. In the tradition RDMS consistency is maintained instantaneously while in NoSQL there may be inconsistency for some time before everything is brought back together.
12. **d) Related data that is often accessed together.** Columns contain related data linked to a key but have similar query issues to key-value
13. **d) All of the above.** RDMS is very rigidly structured and its usage must conform to strict set of principles. Productivity can be improved by using a NoSQL database or system which more closely matches the specific needs. Access to data and scalability can be improved by storing the information on many nodes which distributes traffic and entry points while also allowing more data to be stored more cheaply.
14. **d) very well suited to problem spaces where we have** Graph databases are very good at mapping relationships and allow for very quick network searches.
15. **b) generally useful for content management systems,** Document databases are great for variable data since almost anything can be stored in them. They also are really good for searching the data stored on them since anything can be queried. Files are simply stored in a hierarchy without any other object-relation information needed.
16. **a) generally useful for storing session information**, Key value databases are very good for unknown or unstructured data but there is no way to query them based on the content of the data, a key must be used.
17. **c) generally useful for content management systems,** Column family databases are also very good for unstructured data that are related to each other but your query structure must be well known and established.
18. **b) Data may be read according to the implicit schema** This assumes that the people writing the software for the database have written their code anticipating possible schema changes.