

Wrap-Up

























Summary: Why Use a DBMS?

- DBMS: serving large data sets to large, heterogeneous user groups
- Quality of service
 - (optimized) query language = fast & flexible access to large data assets
 - Concurrent access
 - Data independence
- Efficiency
 - scalability; reduced application development time
- Information integration
 - Uniform data administration; concise information modelling
- Safety
 - Data integrity & security; Crash recovery

Big Data

DB Ranking by Deployments

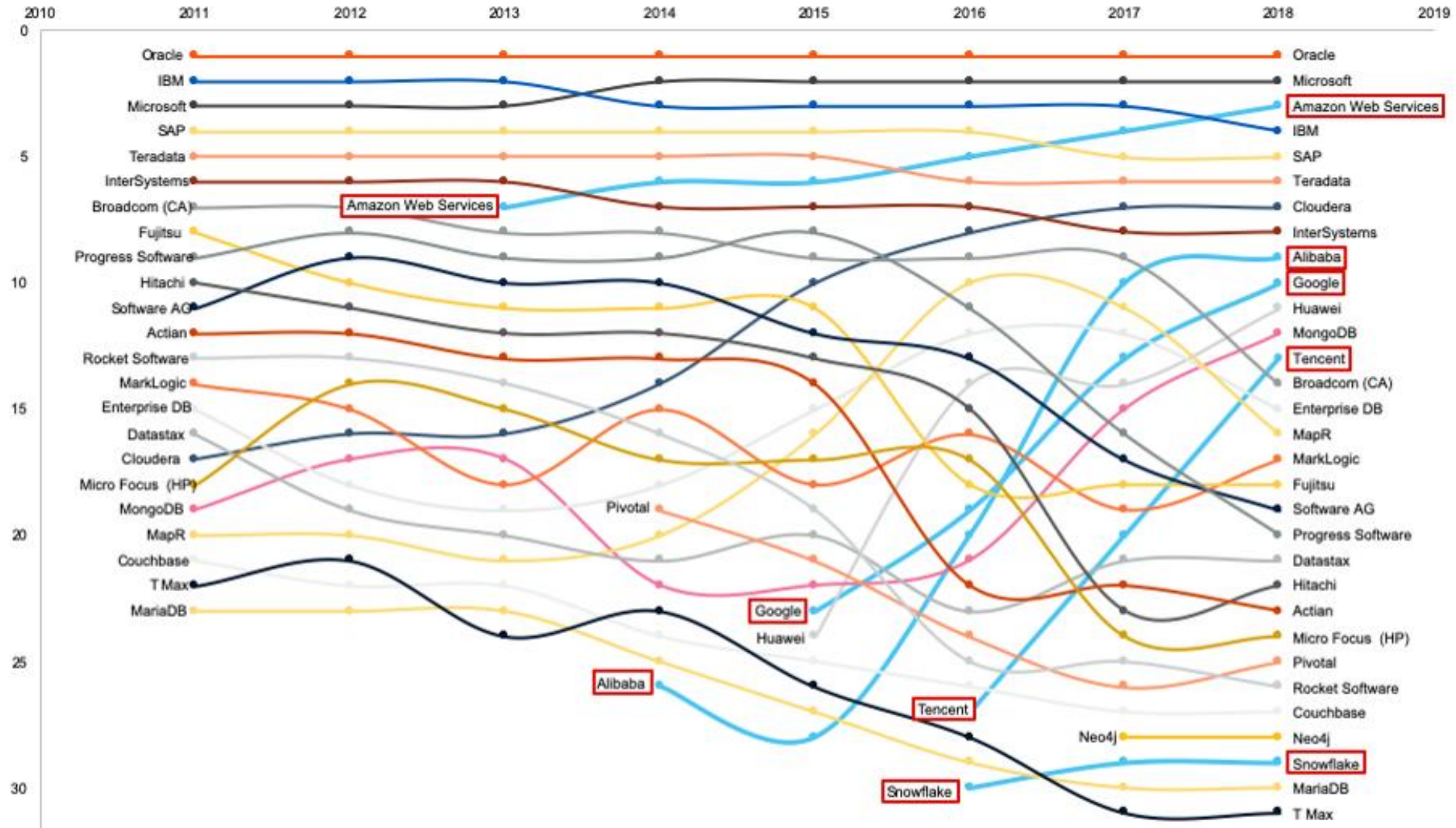
356 systems in ranking, June 2020

Rank			DBMS	Database Model	Score		
Jun 2020	May 2020	Jun 2019			Jun 2020	May 2020	Jun 2019
1.	1.	1.	Oracle 	Relational, Multi-model 	1343.59	-1.85	+44.37
2.	2.	2.	MySQL 	Relational, Multi-model 	1277.89	-4.75	+54.26
3.	3.	3.	Microsoft SQL Server 	Relational, Multi-model 	1067.31	-10.99	-20.45
4.	4.	4.	PostgreSQL 	Relational, Multi-model 	522.99	+8.19	+46.36
5.	5.	5.	MongoDB 	Document, Multi-model 	437.08	-1.92	+33.17
6.	6.	6.	IBM Db2 	Relational, Multi-model 	161.81	-0.83	-10.39
7.	7.	7.	Elasticsearch 	Search engine, Multi-model 	149.69	+0.56	+0.86
8.	8.	8.	Redis 	Key-value, Multi-model 	145.64	+2.17	-0.48
9.	9.	 11.	SQLite 	Relational	124.82	+1.78	-0.07
10.	 11.	10.	Cassandra 	Wide column	119.01	-0.15	-6.17
11.	 10.	 9.	Microsoft Access	Relational	117.18	-2.72	-23.83
12.	12.	12.	MariaDB 	Relational, Multi-model 	89.79	-0.30	+4.59
13.	13.	13.	Splunk	Search engine	88.08	+0.33	+3.46
14.	14.	14.	Hive	Relational	78.65	-2.89	-0.40

DB Ranking by Dev'ers, Stackoverflow

1. MySQL
2. PostgreSQL
3. Microsoft SQL Server
4. SQLite
5. MongoDB
6. Redis
7. MariaDB
8. Oracle
9. Firebase
10. Elasticsearch

Gartner Market Share Ranking



„No One Size Fits All“

- General insight today: **no singular data modeling paradigm** (eg, sets) can match all requirements in semantics & performance
- Ex: SAP HANA: four main-memory storage engines
 - column-store, for OLAP-dominant & mixed workloads
 - row-store, for OLTP-dominant workloads
 - graph engine
 - text engine



<http://martinfowler.com/bliki/PolyglotPersistence.html>

■ Refs:

- M. Stonebraker et al. *"One size fits all": an idea whose time has come and gone*. ICDE, 2005
- F. Färber et al.: *The SAP HANA Database – An Architecture Overview*. IEEE Data Eng. Bull., 35(1):28–33, 2012
- V. Sikka et al. *Efficient transaction processing in SAP HANA database: the end of a column store myth*. SIGMOD, 2012

Course Plot – or: why did I take it?

- How to design databases, and how to search them
- How to design (Internet) services

- Database services revisited
- Practice: set up a Web service
 - LAMP = Linux, Apache, MySQL, PHP

What industry expects a CS graduate to know

Your entry point to the DB [admin] world

Check also database videos, such as [this one](#)

Must-Haves for IT Job Interviews

- "47% of the job ads analyzed expect **economics knowledge**. Also, **communication skills** are emphasized.
- Currently **database skills** are at the top of the IT companies' wish list, every 3rd IT job ad requires them. Further, Business Intelligence, Enterprise Resource Planning, and Service-Oriented Architectures are an asset. Additionally, relevant hands-on experience, e.g., in project work, plays an important role."
 - Thomas Reher, Executive Board member, PPI AG

But, Mind You:

Best Success!

- Chances are you won't use classroom knowledge as is
 - Diversity of technology, requirements, enterprise setups, ...
- ...then why did we do it??
 - **Grasp the concepts**
 - Whatever gossip says – SQL is like English: y'all just need to know (at least basics)
 - Able to immerse into any DB & Web services technology rapidly

