## 1. Cloud computing

Alright, this is the final stretch!

## 2. Cloud computing for data processing

Let's now talk about cloud computing. Companies can process data in their own data center, often on premises. We can imagine racks of servers, ready to be used, that the company has to buy. We also need a room to store them, and if we move offices, we have to transport servers without losing service. The electrical bill and maintenance would be at the company's cost. Moreover, data processing tasks can be more or less intense, and don't happen continuously. Companies would need to provide enough processing power for peak moments, and at quieter times, much of the processing power would remain unused. It's avoiding this waste of resources that makes cloud computing so appealing. In the cloud, we rent servers, and the rent is cheap: we don't get the same discount from our server salesperson that Amazon). We don't need a room to store them, and we use the resources we need, at the time we need them. Many companies moved to the cloud as a way of cost optimization. Also, the closer the server is to the user, the less latency they will experience when using our application. To serve a global customer base, we need servers all over the world.

## 3. Cloud computing for data storage

Another reason for using cloud computing is database reliability. Running a data-critical company, we have to prepare for the worst. A fire can break out in an on-premises data center. To be safe, we need to replicate our data at a different geographical location. However, if your company manipulates sensitive or confidential data, there is a risk associated with someone else hosting it, and government surveillance. With the advent of big data, companies specializing in these kinds of issues, called cloud providers, were born.

## 4. Cloud providers

The three big players, in decreasing order of market share, are Amazon Web Services,

## 5. Cloud providers

Microsoft Azure,

## 6. Cloud providers

and Google Cloud.

## 7. Cloud providers

For file storage,

## 8. Cloud providers

their respective services are AWS S3,

## 9. Cloud providers

Azure Blob Storage,

## 10. Cloud providers

and Google Cloud Storage.

## 11. Cloud providers

For computation,

## 12. Cloud providers

they are AWS EC2,

## 13. Cloud providers

Azure Virtual Machines,

## 14. Cloud providers

and Google Compute Engine.

## 15. Cloud providers

For databases,

## 16. Cloud providers

they are AWS RDS,

## 17. Cloud providers

Azure SQL Database,

## 18. Cloud providers

Google Cloud SQL.

## 19. Cloud computing at Spotflix

Spotflix chose AWS,

## 20. Cloud computing at Spotflix

so we use S3 to store cover albums,

## 21. Cloud computing at Spotflix

EC2 to process songs,

## 22. Cloud computing at Spotflix

and RDS to store employees information.

## 23. Multicloud

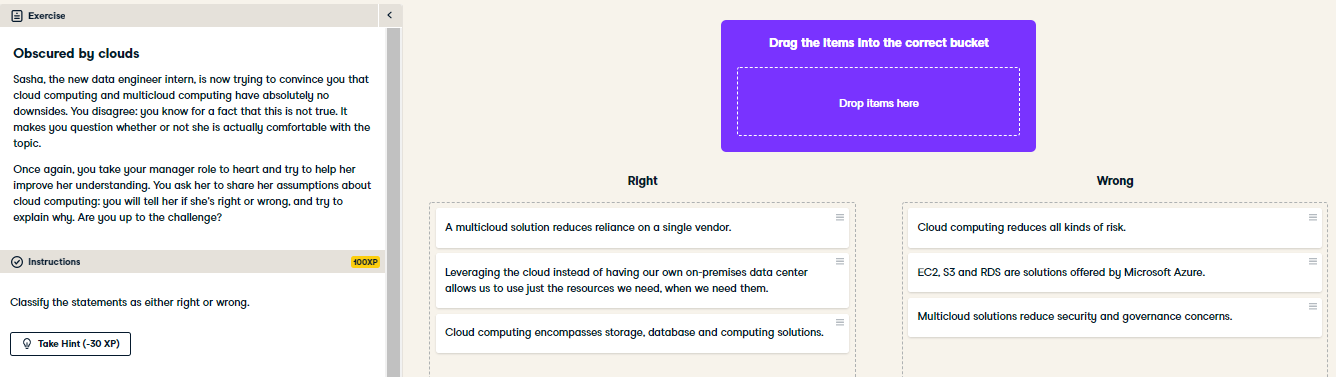
You don't need to take all your cloud services from the same provider though. You can use several ones: it's called multicloud. It has some advantages, like reducing reliance a single vendor. It also optimizes costs, and might be necessary because of local laws. It also allows you to militate against disasters. For example, in 2017, AWS had an outage, which broke the internet: among other companies impacted half of the top hundred retailers were down. I mean, even isitdownrightnow dot com was down. Spotflix wasn't impacted in the end, because, well, it doesn't exist. Companies that had some redundancy with other providers could mitigate the impact. However, cloud providers try to lock in consumers, by integrating as many of their services as they can. Some services from one provider may not be compatible with services from another one, so that's something to be careful about. It also makes managing security and governance harder.

## 24. Summary

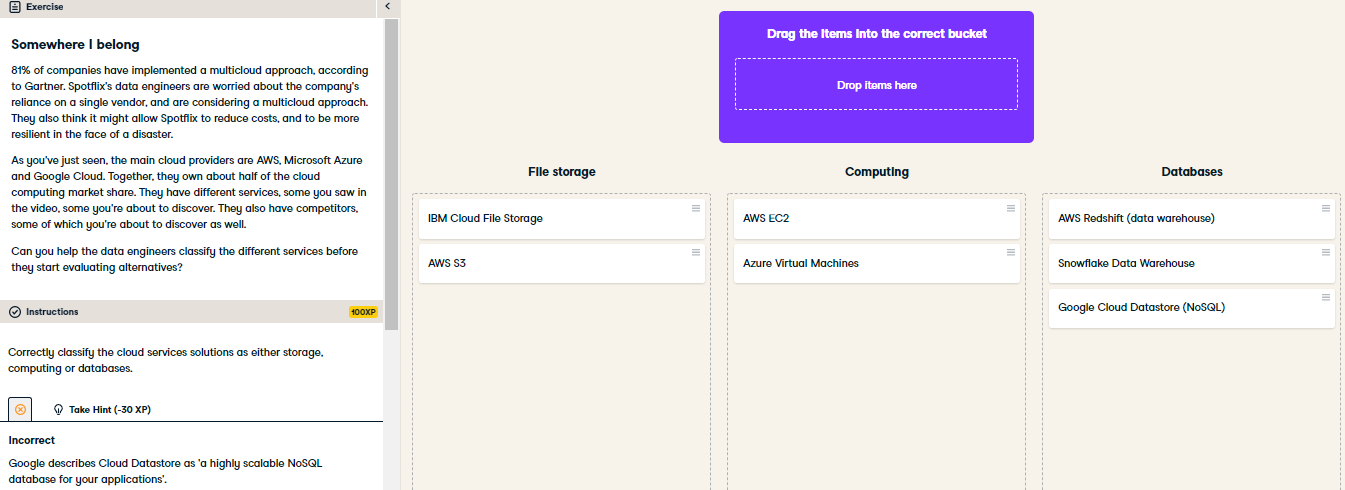
Alright, now you know the benefits and risks of cloud computing, and how it is implemented at Spotflix. You can also cite the main cloud providers and their respective services.

## 25. Let's practice!

Home stretch with some exercises, and then I will tell you where to find the playlist, as promised at the end of Chapter 1!



Well played! You've empowered your data engineer intern. Now, not only does she have a more nuanced point of view and knows she can come to you with questions, you also sense that your intern is better prepared for her job.



Snowflake, founded in 2012, is disrupting the data warehouse industry and enjoyed a 174% revenue growth in 2019. That might be a name to remember!