

And the data schema and evolution, so how the data is shaped and how it may change in the future. Additionally, each time you integrate a source system with a target system, there will be an increased load from the connections. So how do we solve this? Well, this is where Apache Kafka comes in. Apache Kafka allows you to decouple your data streams and your systems. So now, your source systems will have their data end up in Apache Kafka while your target systems will source their data straight from Apache Kafka. This decoupling is what is so good about Apache Kafka and what it enables is really, really nice. For example, what do we have in Kafka? Well, you can have any data stream you can think about. For example, it could be website events, pricing data, financial transaction, user interactions and many more. Additionally, once the data is in Kafka, you may want to put it into any system you like, such as a database, your analytic systems, your e-mail system or your audits. Why would we even use Apache Kafka? Well, it was created by LinkedIn and it's now an Open Source project mainly maintained by a private company called Confluent, but it's under the Apache stewardship. It's distributed, resilient architecture and fault-tolerant and basically, it scales. It scales horizontally. There are Kafka clusters with over 100 brokers. It can also scale and is proven by LinkedIn and many other companies that you can scale to millions of messages exchanged per second. That's quite a lot of scale. It is extremely high performance. The latency to exchange data from one system to another is usually less than 10 millisecond if you have good machines. This is what we call real-time. Real-time means that the latency is really, really low. Overall, don't believe me, right? You can believe 2,000 plus firms, 35% percent of the Fortune 500 that use Kafka, Such LinkedIn, Airbnb, Netflix, Uber, Walmart and so many others. Everyone uses Kafka or is thinking about using Kafka and you should learn Kafka and you are at the right place. Use cases for Apache Kafka. Well, it could be a messaging system, it could be used for activity tracking or to gather metrics from many different locations, (mumbles) your examples or your IoT devices. Gather logs from your applications, stream processing, and we'll see that we can use the Kafka Streams API for that. Decoupling of system dependencies, so reduce the load on your databases and your systems by decoupling them. Or perform big data integrations, for example with Spark, Flink, Storm, Hadoop and other big data technologies. This is quite a wide array of use cases, and this is why so many companies are using Apache Kafka as their backbone in their company. Now, if I wanted to get you concrete example, get you a bit excited about Kafka, that make you understand what Kafka is used for; Netflix is using Kafka to apply recommendations in real-time while you're watching TV shows. And this is why basically, when you leave a TV show, you'll get a new recommendation right away. Uber uses Kafka to gather user, taxi and trip data in real-time to compute and forecast demand and compute the almighty surge pricing in real-time. So Uber uses Kafka extensively. LinkedIn uses Kafka to prevent spam in their platform, collect user interactions and make better connection recommendations. All of that in real-time. Basically, as you can see, all these companies are using Kafka so that they can make real-time recommendation, real time decisions, give you real-time insights to their users and this is why it's so good. Now, remember that in this, Kafka is only used as a transportation mechanism. People initially still write their applications or web applications to make things work, but Kafka is really

good at making your data move really fast, at scale in your company. So that's it for a quick introduction, to give you some insight about what Kafka is, how it works and where it sits in companies and how they use them. In the next lectures and in the entire course, you're going to accumulate so much knowledge and practice around Kafka. But I hope that in these first five minutes, you really understand what we're going to talk about for the rest of this course and I hope you're excited. I will see you in the next lecture.

Give feedback