# Explain how cloud architecture works with Data Engineering

To explore how cloud architecture works with data engineering, it’d be helpful to compare data engineering before cloud architecture and data engineering after cloud architecture. My previous employer’s main product is a SaaS solution that appears to be run “on the cloud” to clients. However, because the company began before cloud computing became widely leveraged, the backend of the solution lives on data centers built and managed by the company itself. The software is still “on the cloud” in the sense that all computations take place on remote servers, but the servers lack many of the capabilities offered by major cloud service providers today – it was quasi-cloud-based.

Since it’s almost impossible to find data engineering teams that are completely disconnected from the cloud today, we can get a glimpse of what data engineering used to look like by examining the tasks of data engineers at my previous employer. Setting up servers, creating databases and tables, and ensuring the health of the servers and databases are major responsibilities of the company’s data engineers. Since the company has its own data centers, the data engineers have to take care of the servers and databases themselves. They also use tools and scripts created by other data engineers likely quite a few years ago to perform such tasks. Compare that to doing similar tasks on the cloud. Service providers already took care of the hard jobs in creating and maintaining servers and databases, so data engineers would only need to write scripts to provisioning servers and databases to the needs of the business, and not have to worry about things like server health.

The data engineers at my previous employer also spent a lot of time creating ETL tools. While creating ETL tools is the textbook definition of what a data engineer does, the way that the ETL tools were created and applied was not the most efficient. My previous employer received data from different clients in vastly different formats, so its data engineers transforming the data before sending it to the data ingestion pipeline. Relying on the old architecture, this often resulted in creating ETL processes from scratch. Before I left, the company started an initiative to templatize the ETL process in order to save data engineering time, but it required a lot of work. On the other hand, using cloud architecture, building templates is much easier with tools like cloud-based docker containers.

Another area that data engineers at my previous employer spent a lot of time on is security. They had built many tools to ensure that the data hosted on the company’s servers stays encrypted and secure. At times, efficiency was greatly sacrificed to ensure security. With major cloud service providers though, security lies at the heart of the service, so data engineers can spend much less time creating methods to keep data secure, and they can spend time elsewhere.

To summarize, cloud architecture makes the work of data engineers more efficient by providing abstractions of hardware, offering containers and other serverless tools, and being secure.