Summary

Thank you for joining us in this module on XCOMs

1. We learnt about XCOMs, a native feature of Airflow, that enables the sharing of data between tasks by using the Airflow meta database.
2. To use XCOM, a DAG and task must first be created, and then an XCOM can be created by returning a value from a task that can be retrieved by another task.
3. The XCOM has a key, value, timestamp, task ID, and DAG ID, and must be JSON serializable.
4. Using the xcom_push and xcom_pull methods in Airflow, one can define a specific key to an XCOM, giving flexibility in defining task IDs and allowing for pulling data from multiple tasks simultaneously.
5. The Airflow UI can display XCOMs and their properties, and XCOMs can be used to share small amounts of metadata between tasks.
6. However, XCOMs have limitations based on the database used, with SQLite allowing up to 2GB, Postgres up to 1GB, and MySQL up to 64KB.
7. It is not suitable for sharing large amounts of data, and for that, one should trigger a Spark job or similar.
8. In summary, XCOMs are a useful feature of Apache Airflow for sharing small amounts of data between tasks, but it is important to understand their limitations and use them accordingly.

Thank you for tuning in. See you in the next one.