

2) Sometimes real time data is needed for analysis. Nevertheless, this should be planned carefully in order to maintain system performance. Design a data pipeline strategy for real time data streams that allows scalability.

For this case I propose to use a ELT (Extract, Load, Transform), the reason of that is ELT allows scalability, because the information is load direct in the server and not in a provisional server, for this reason the load and transformation can occur at the same time

Data Collection

