

## **How do data stream systems differ from traditional relational databases?**

Fundamental difference: data stream model.

In a data stream, data elements arrive on-line and stay only for a limited time period in memory. Consequently, the DSMS has to handle the data elements before the buffer is overwritten by new incoming data elements. The size of data streams is potentially unbounded and can be thought of as an open-ended relation.

## **Mention a couple of data stream applications, and explain why a RDBMS would not be able to support them.**

Financial real-time analysis; video streaming; network monitoring and traffic engineering; web logs and click-streams. For those applications, input data that comes at a very high rate. High rate means it stresses communication and computing infrastructure, so it may be hard to

- transmit (T) the entire input to the program,
- compute (C) sophisticated functions on large pieces of the input at the rate it is presented, and
- store (S), capture temporarily or archive all of it long term.

Thus in a RDBMS, it's hard to computer and store those high-rate and unbounded data.

**Why do you think that exact query processing is very hard to achieve in the context of data streaming?**

Because a data stream is a real-time, continuous, ordered(implicitly by arrival time of explicitly by timestamp) sequence of items. It is impossible to control the order in which items arrive, not it is feasible to locally store a stream in its entirety. But exact query processing acquire the entirety of data.