Why is Streaming Data Different?

After this video you will be able to..

- Compare and contrast “data-in-motion” and “data-at-rest”
- Differentiate between streaming and batch data processing
- List management and processing challenges for streaming data

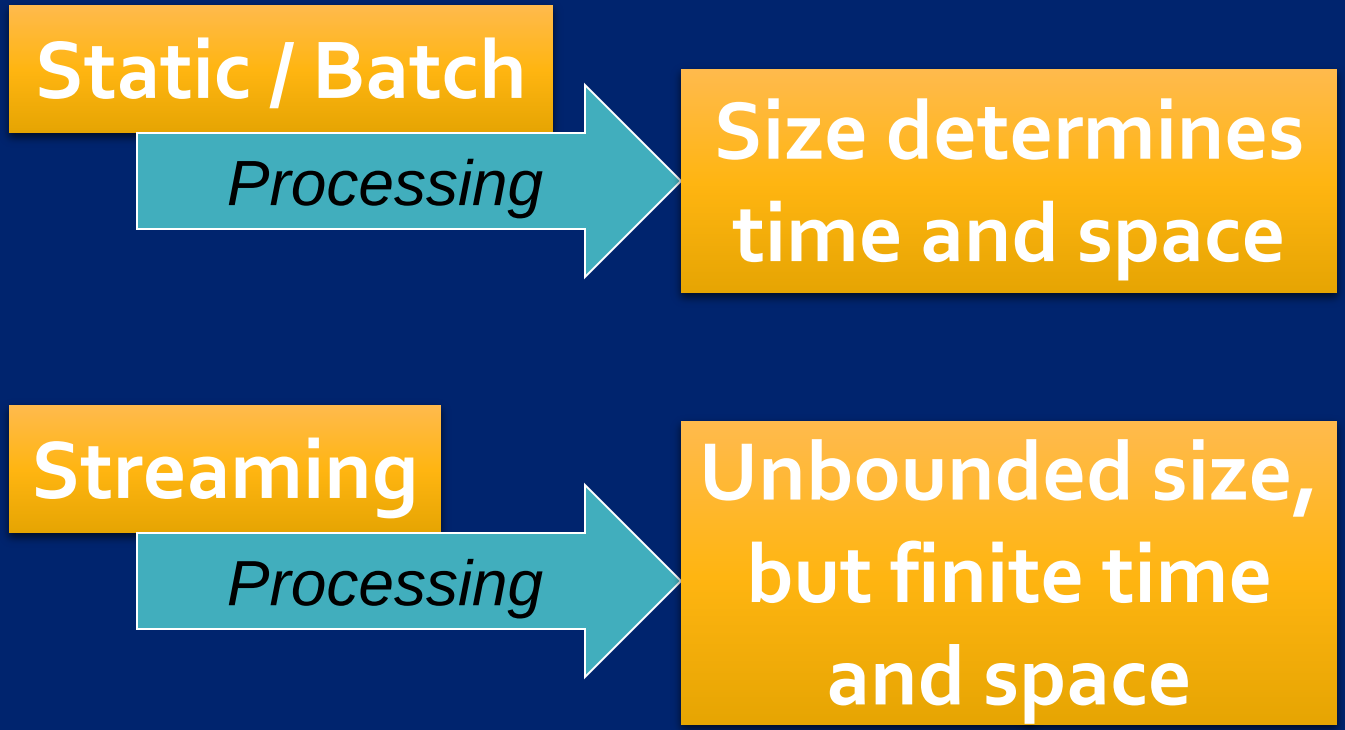
Data-at-Rest

- Mostly static data from one or more sources
- Collected prior to analysis

Data-in-Motion

- **Analyzed as it is generated**
  - Example: sensor data from self-driving vehicles
- **Stream processing**

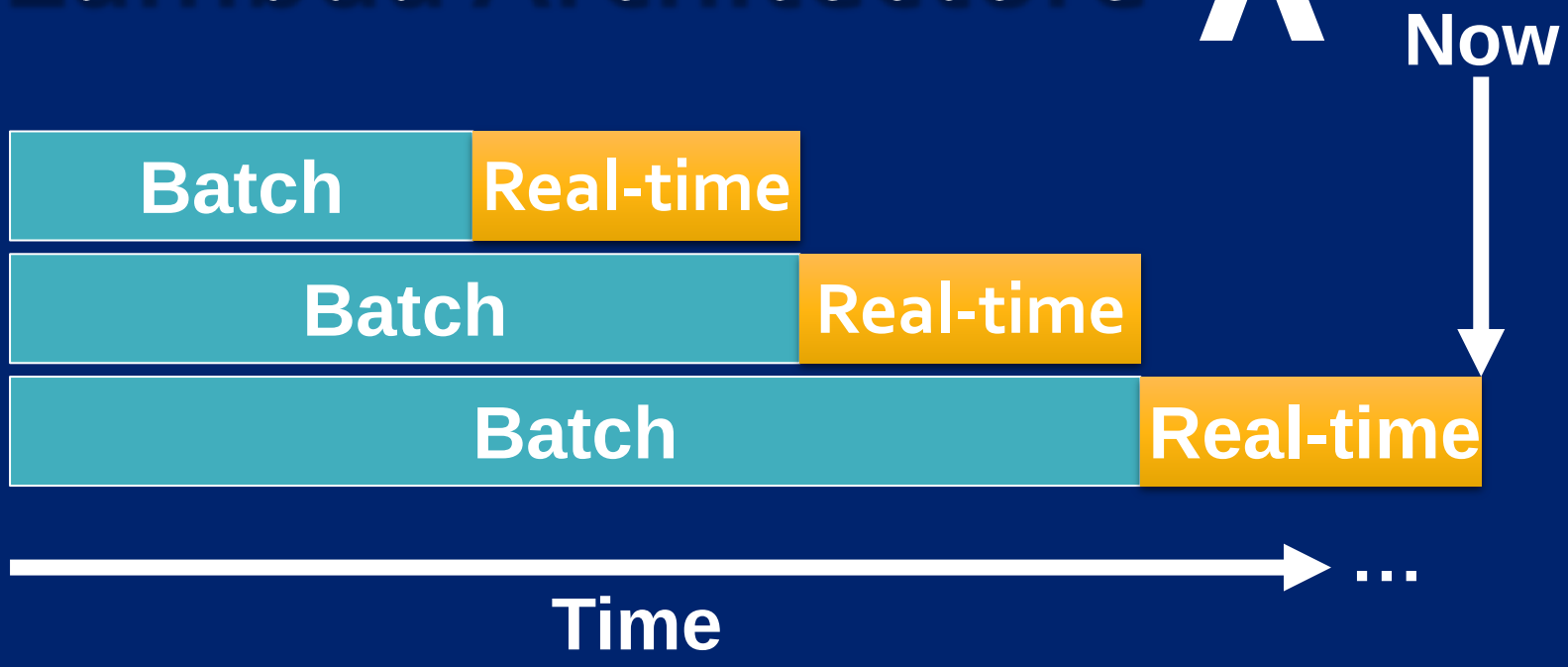
Data Processing Algorithms



Streaming Data Management and Processing

- Compute one data element or a small window of data elements at a time
- Relatively fast and simple computations
- No interaction with the data source

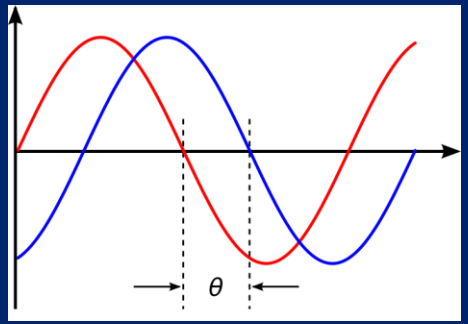
Lambda Architecture



Streaming Data Changes Over Time

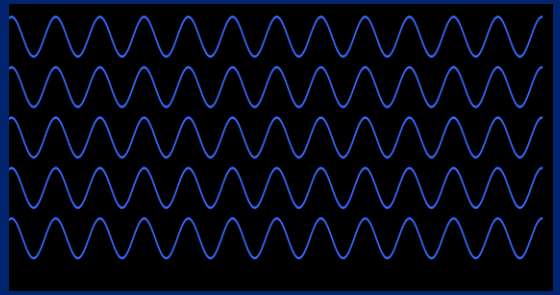
Size

Frequency



Changes can be periodic or sporadic

Periodic: evenings, weekends, etc.



Sporadic: major events

Example of extreme change:
World Record for Tweets

Average Tweets / Second	= 6000
Record Tweets / Second	> 144,000

Streaming Data Summary

- Size → Unbounded
- Size and Frequency →Unpredictable
- Processing → Fast and Simple