Section 4: Week 7: Time Series Data Management

Nate Bachmeier

TIM-7020: Database and Business Intelligence

January 19th, 2020

North Central University

Time Series Data Management

## Background

Data volumes continue to grow exponentially on an annual basis because of IoT, Cloud, Big Data, and Mobile (ICBM). This massive emersion of records introduces challenges for organizations as they need to gain timely insights into their data. Consider the raw resource costs associated with the data management lifecycle in terms of computing, storage, and networking. For instance, the Hadron collider generates 300GB/s, making an hour of sensor data approximately 8.6Pb (Basanta-Val et al., 2017)! Other examples exist across health care (e.g., medical monitoring devices), finance (e.g., high-frequency trading), retail (e.g., click streams), and manufacturing (e.g., industrial IoT) sectors as users and devices produce these time-series data streams. Traditional big data systems addressed these issues by staging data into an object or NoSQL store, then executing batch processing models such as Map-Reduce (Barika & Garg, 2019). While this approach works today, it will continue to become economically prohibitive going forward. Instead, the programming models need to evolve towards real-time stream processing that extracts and store subsets from the feed. As the programming model changes to iterative processing paradigms, it will cascade across the technology stack, fundamentally changing downstream interactions with data management systems. For instance, training deep learning models relies on offline batch processing versus real-time only systems that will require iterative learning algorithms (Bosch, Olsson, & Brinne, 2019) (Yang et al., 2019). Other changes will take place across infrastructure and business continuity monitoring, as time-series streams are highly contextual—traffic to the eCommerce site is 1000% higher than normal, are we under attack, or did our marketing campaign go viral? There are numerous other scenarios impacted, all of which increase agility for organizations towards dynamic market conditions.

## Problem Statement

## Relevance and Significance

# Literature Review