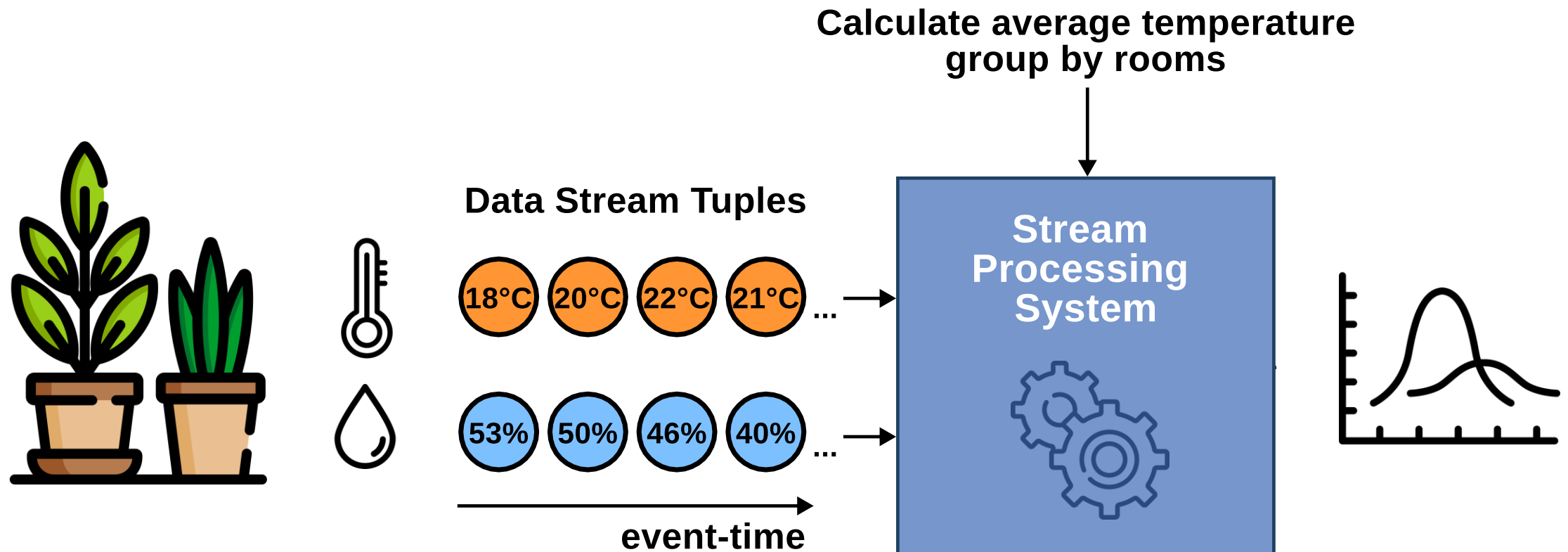


Survey of Window Types for Aggregation in Stream Processing Systems

Juliane Verwiebe, Philipp M. Grulich, Jonas Traub, Volker Markl

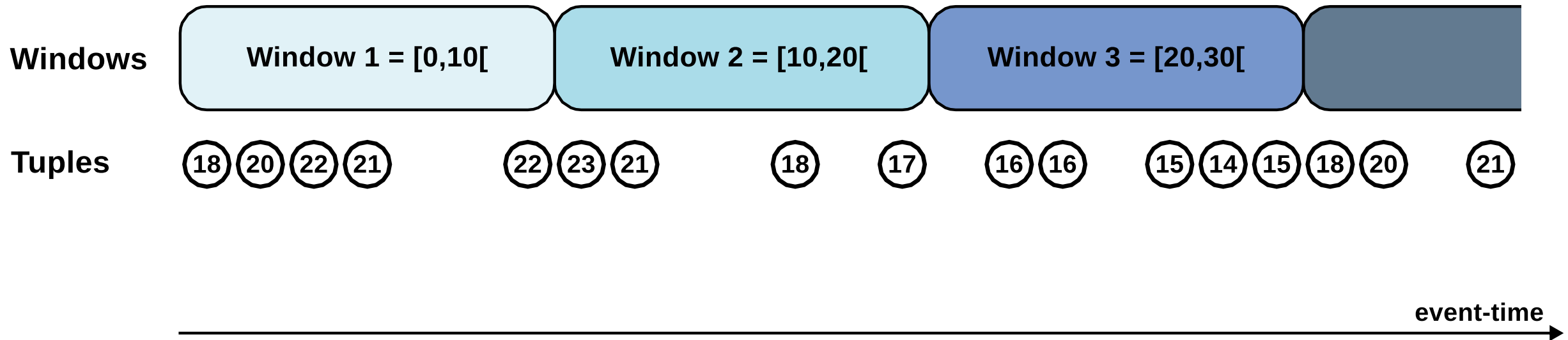
Smart Garden Application



How to calculate aggregates over unbounded streams?

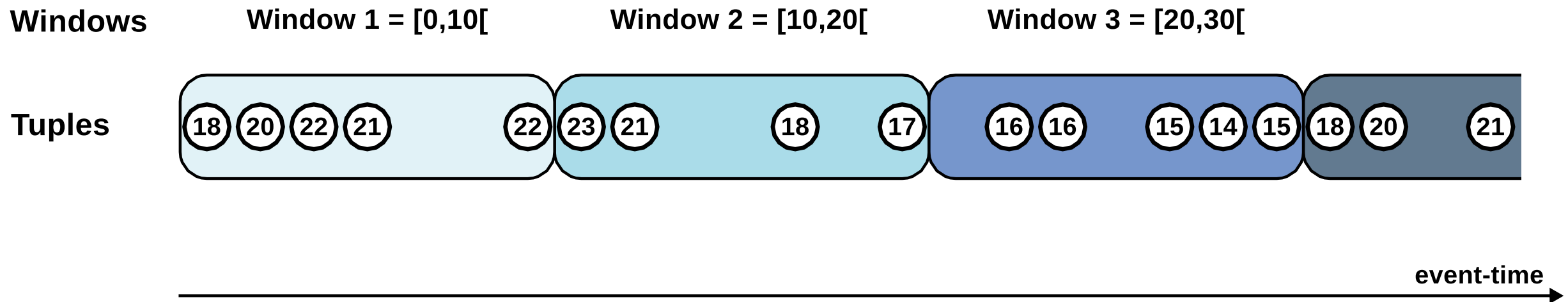
Windows

Divide the data stream into bounded subsets



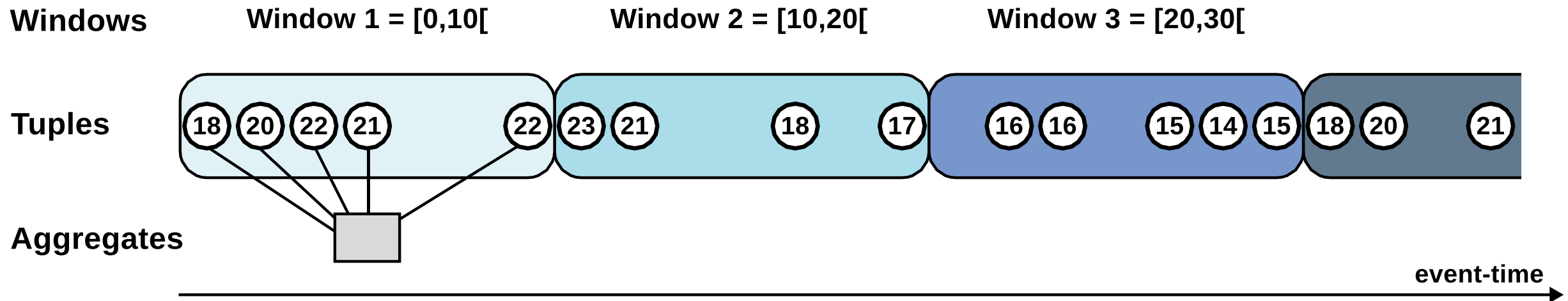
Windows

Divide the data stream into bounded subsets



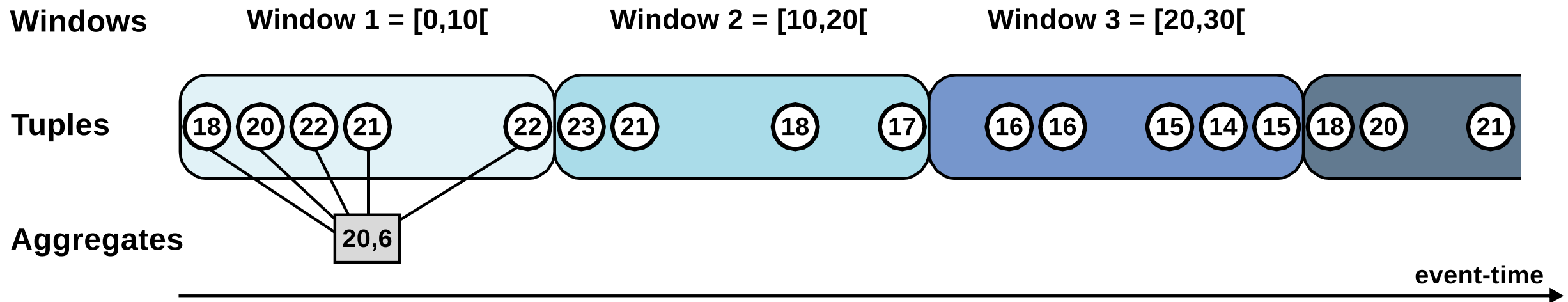
Windows

Divide the data stream into bounded subsets



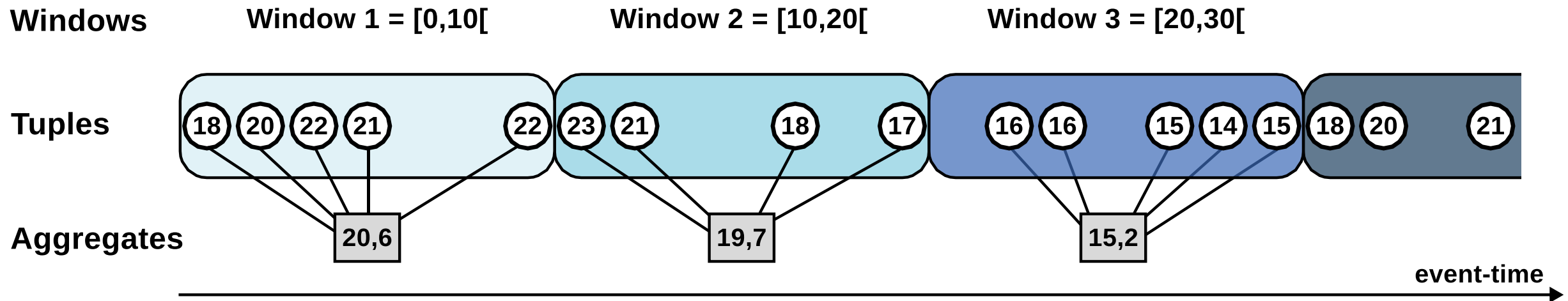
Windows

Average temperature within 10 minutes:



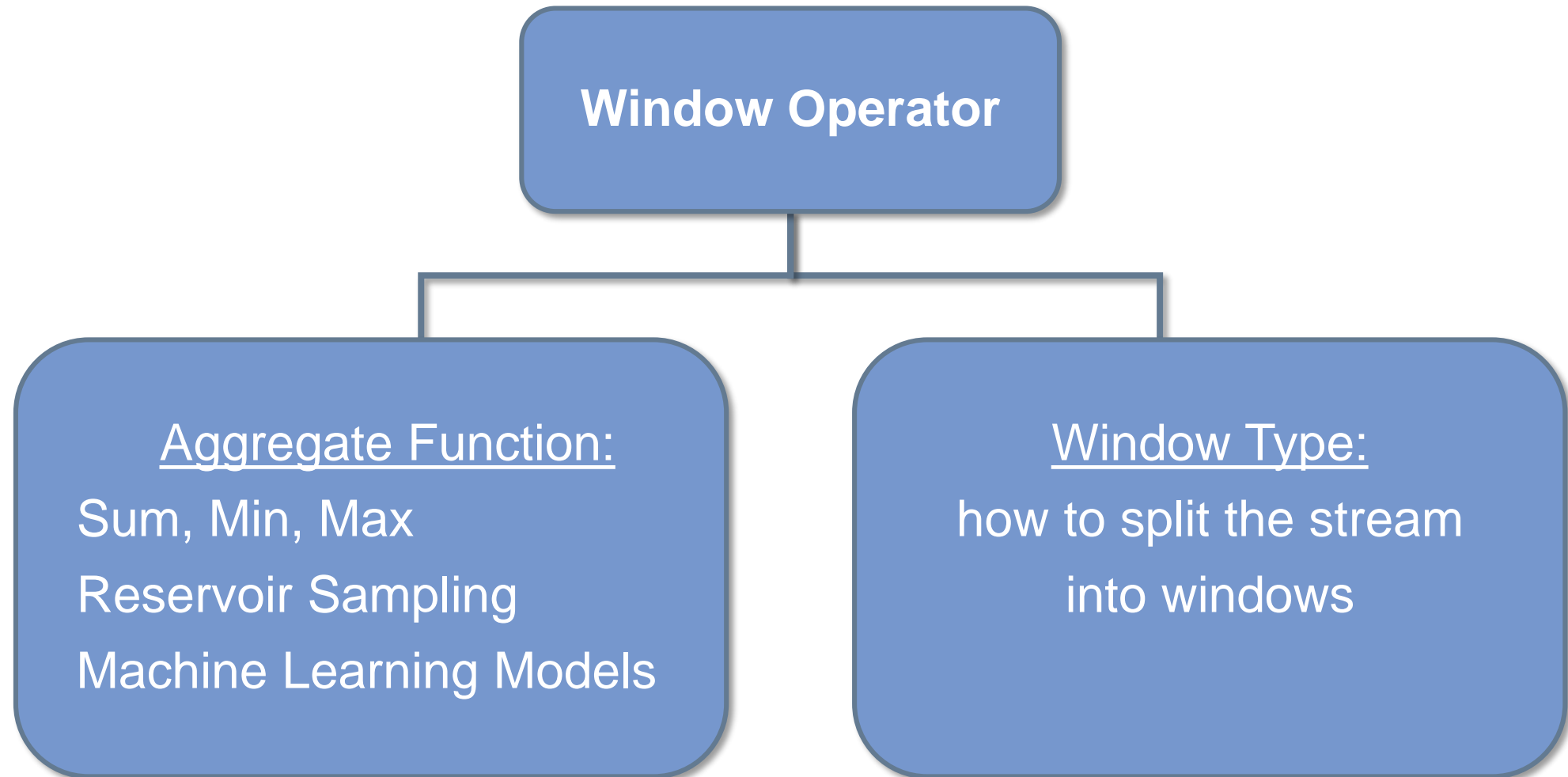
Window Aggregation

Calculate aggregates over windows



Windows enable systems to answer queries

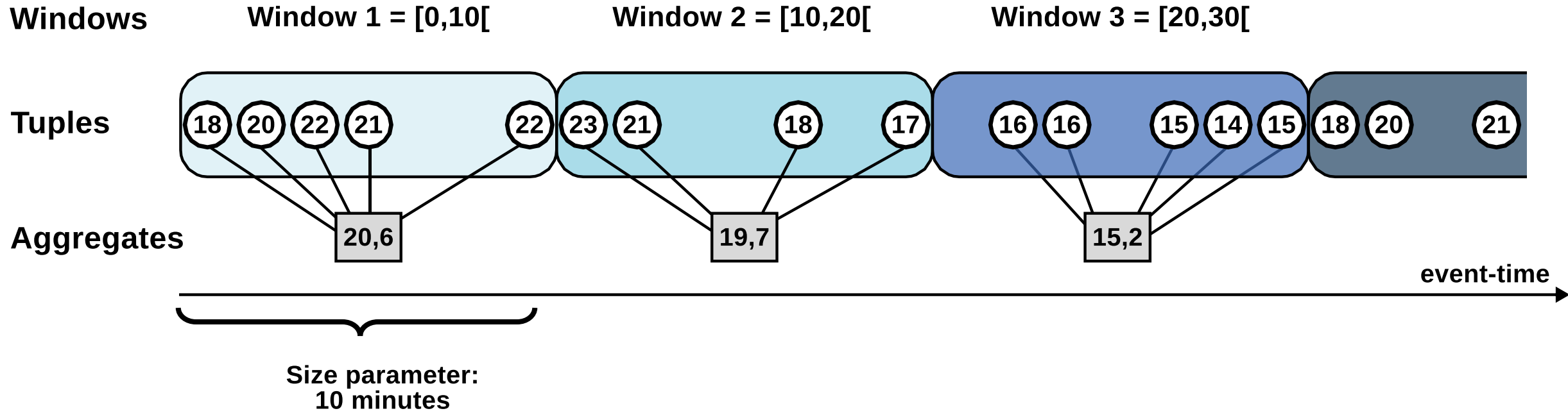
Window Aggregation



Tumbling Window



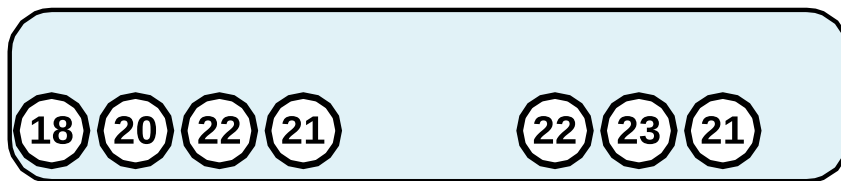
Average temperature in the living room in the last 10 minutes



Sliding Window



Minimum moisture level of the last hour
updated every 5 minutes



Size parameter:
60 minutes

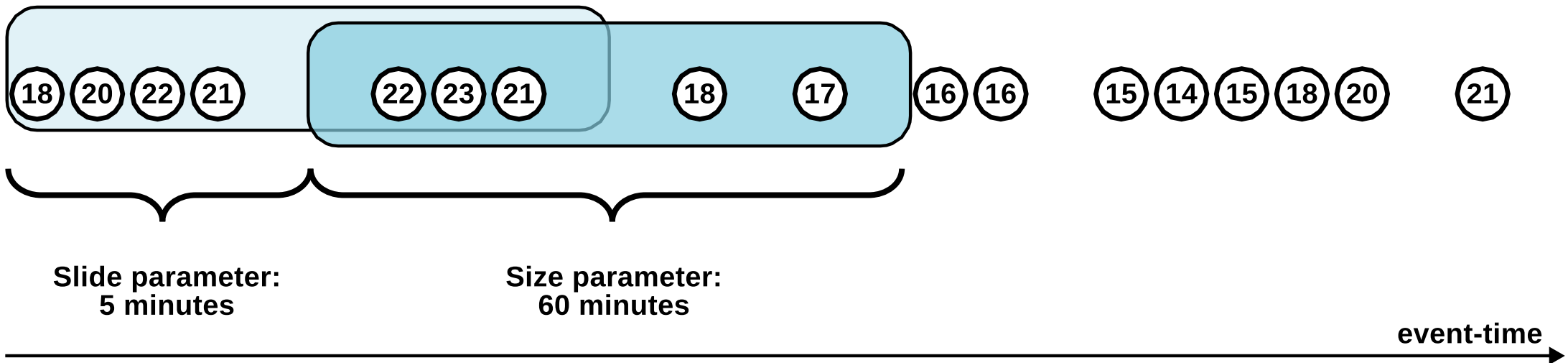


event-time →

Sliding Window



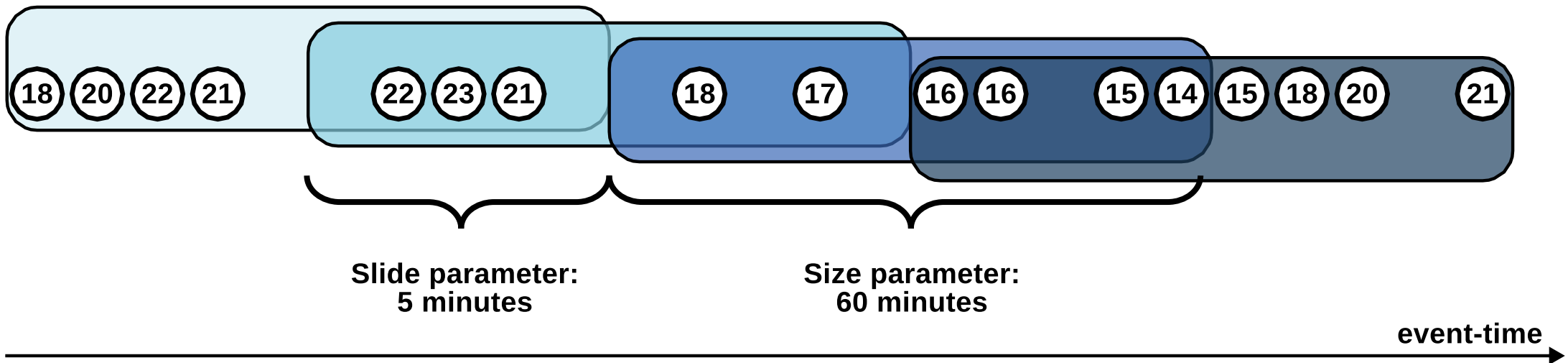
Minimum moisture level of the last hour
updated every 5 minutes



Sliding Window



Minimum moisture level of the last hour
updated every 5 minutes

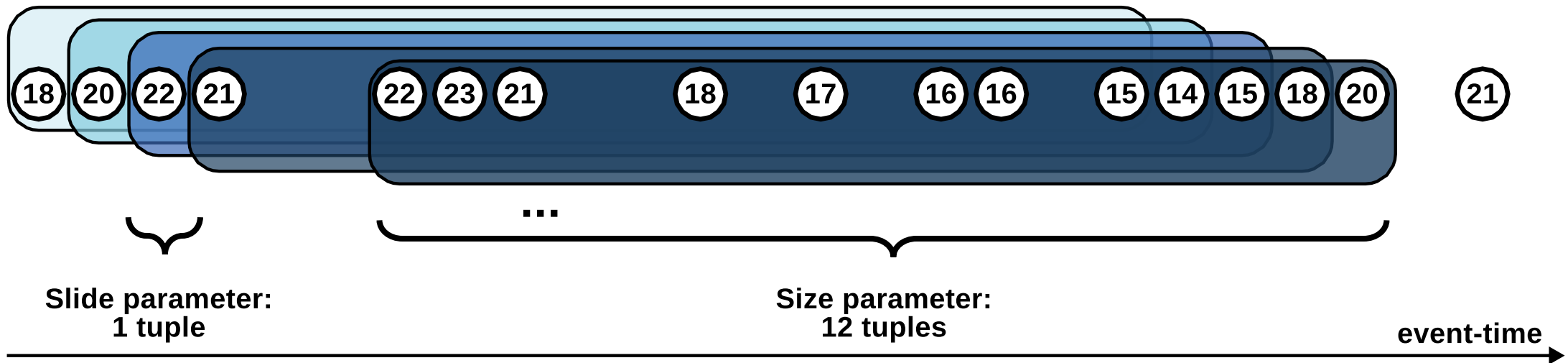


Parameters can be defined using different measures

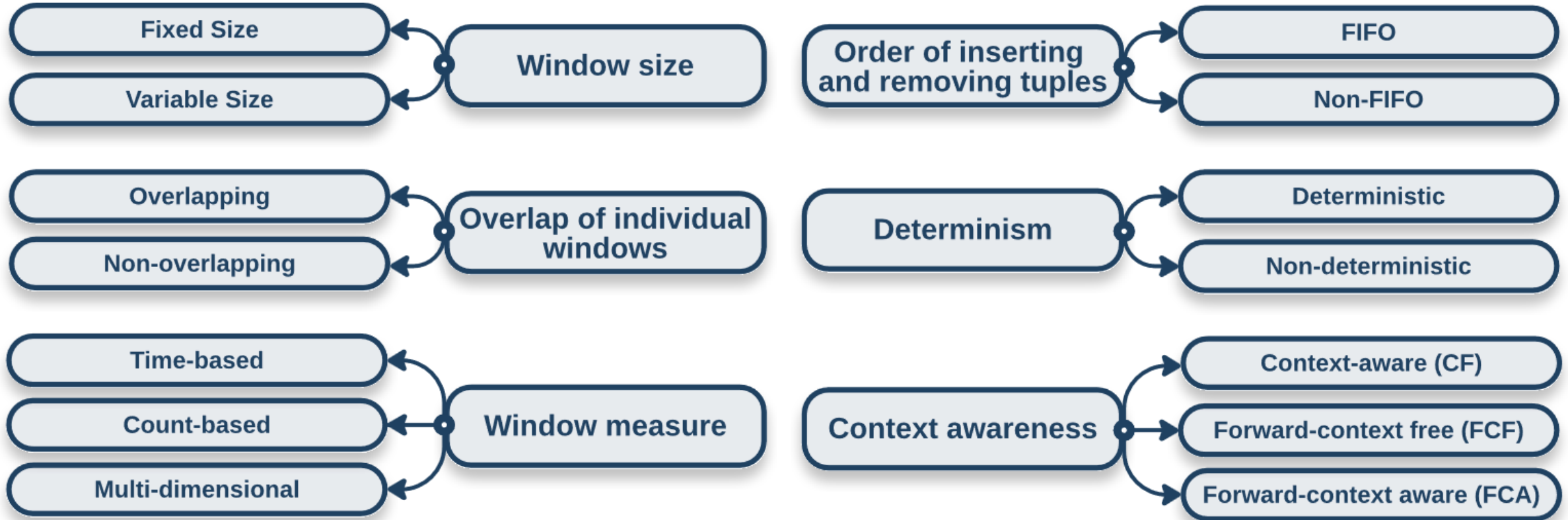
Sliding Window – Count-based



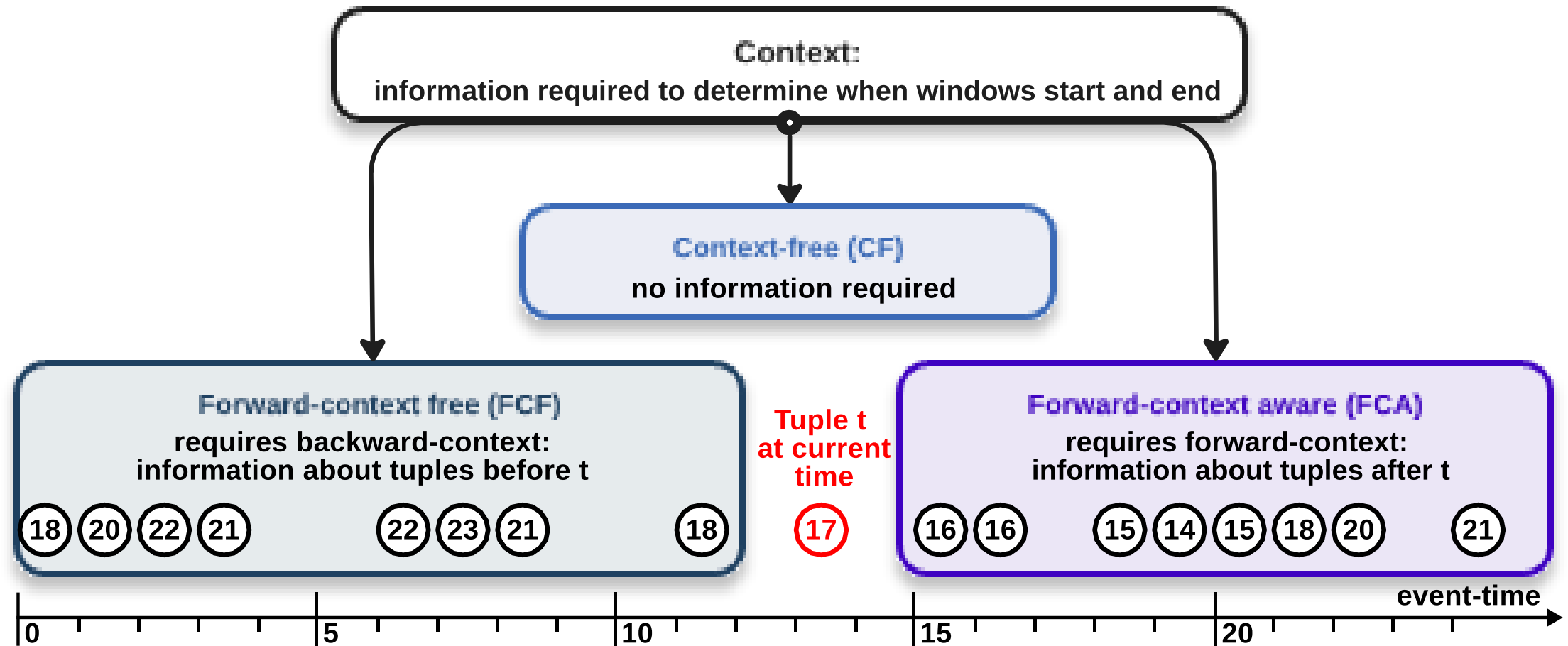
Trend curve: rolling average over measurements



Classification of Window Types



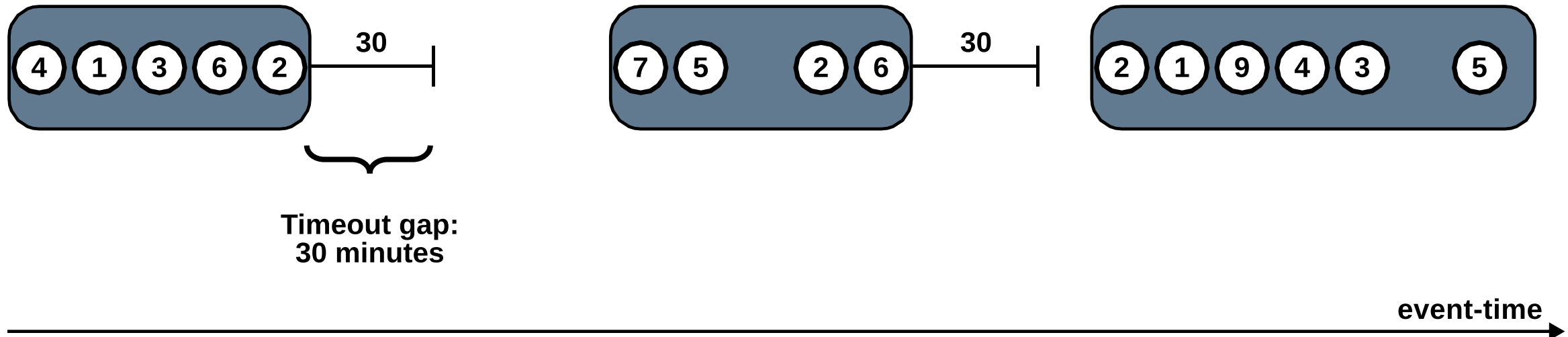
Context-awareness



Forward-context free: Session Window



Minutes of sunlight in a day

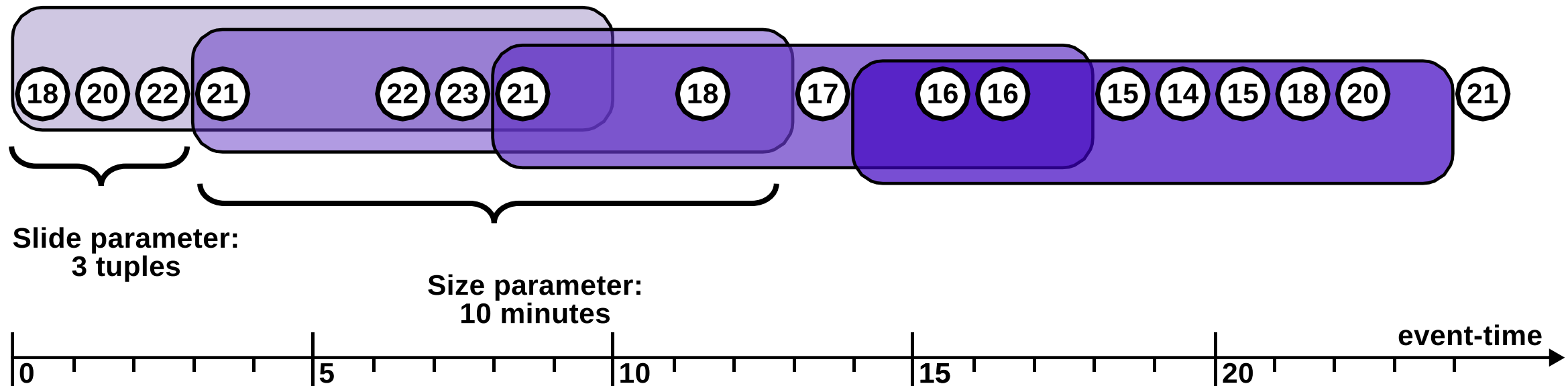


Forward-context aware: Slide-by-tuple Window

Combines time-based size with count-based slide



Update the average
after 3 new measurements



Li et al. 2005

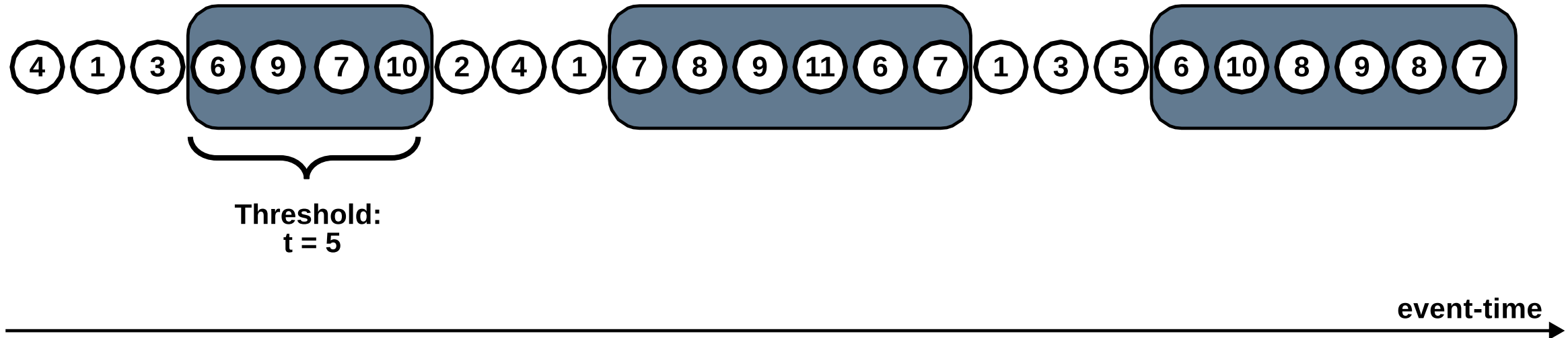
Survey of Window Types

Window Type	Short Description	Window Size	Over-lapping	Window Measure	FIFO	Deter-ministic	Context-awareness
Tumbling Window	Creates periodic, consecutive windows of equal size without gaps	Fixed	No	Time/Count	✓	✓	CF
Sliding Window	Defines window edges by size and slide, slide determines next window start	Fixed	Yes	Time/Count	✓	✓	CF
Session Window	Ends window after timeout gap when no tuples arrive in a predefined time interval	Variable	No	Time	✓	✓	FCF
Slide-by-Tuple Window	Creates a sliding window with size measured in time and slide measured in tuples	Fixed	Yes	Time	✓	X	FCA
...							

16 different window types

Threshold Frame

Value of attribute is above / below a threshold t



Grossniklaus et al. 2016

Stream Processing Systems

Academic

Borealis


STREAM

Telegraph CQ

Naiad

Trill

Commercial

 **amazon**
Kinesis Data
Analytics


Google Cloud
Dataflow

IBM
Streams

 **Microsoft**
Azure Stream
Analytics

Open-source

 **beam**

 **Flink**

 **samza**

 **STORM**

Conclusion

- Overview of window types in literature
- Guideline for the development of systems, frameworks, and window aggregation techniques
- Means to find a suitable window type or system

**Read our Paper
in the VLDB Journal**



**Visit our Website
for open-source
Window Implementations**

