

Survey of Window Types for Aggregation in Stream Processing Systems

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



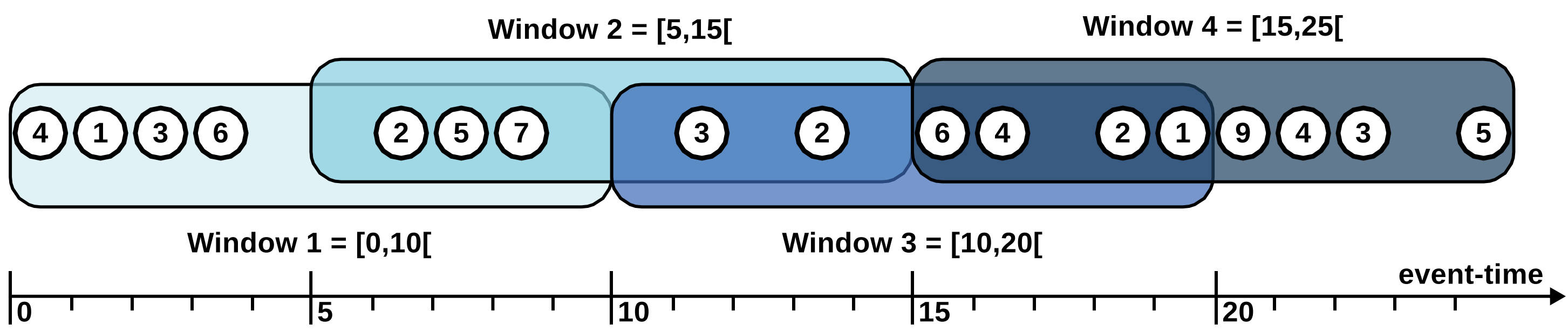
Abstract & Contributions

We present the first comprehensive survey of window types for stream processing systems which have been presented in literature.

- We make the following contributions:
1. We provide a thorough survey of window types.
 2. We classify all window types in categories with respect to performance characteristics and implementation requirements.
 3. For each window type, we provide a formal notation, use-cases, investigate how they can be combined with out-of-order processing, and list synonyms.
 4. We provide an overview of stream processing systems and their support for different window types.

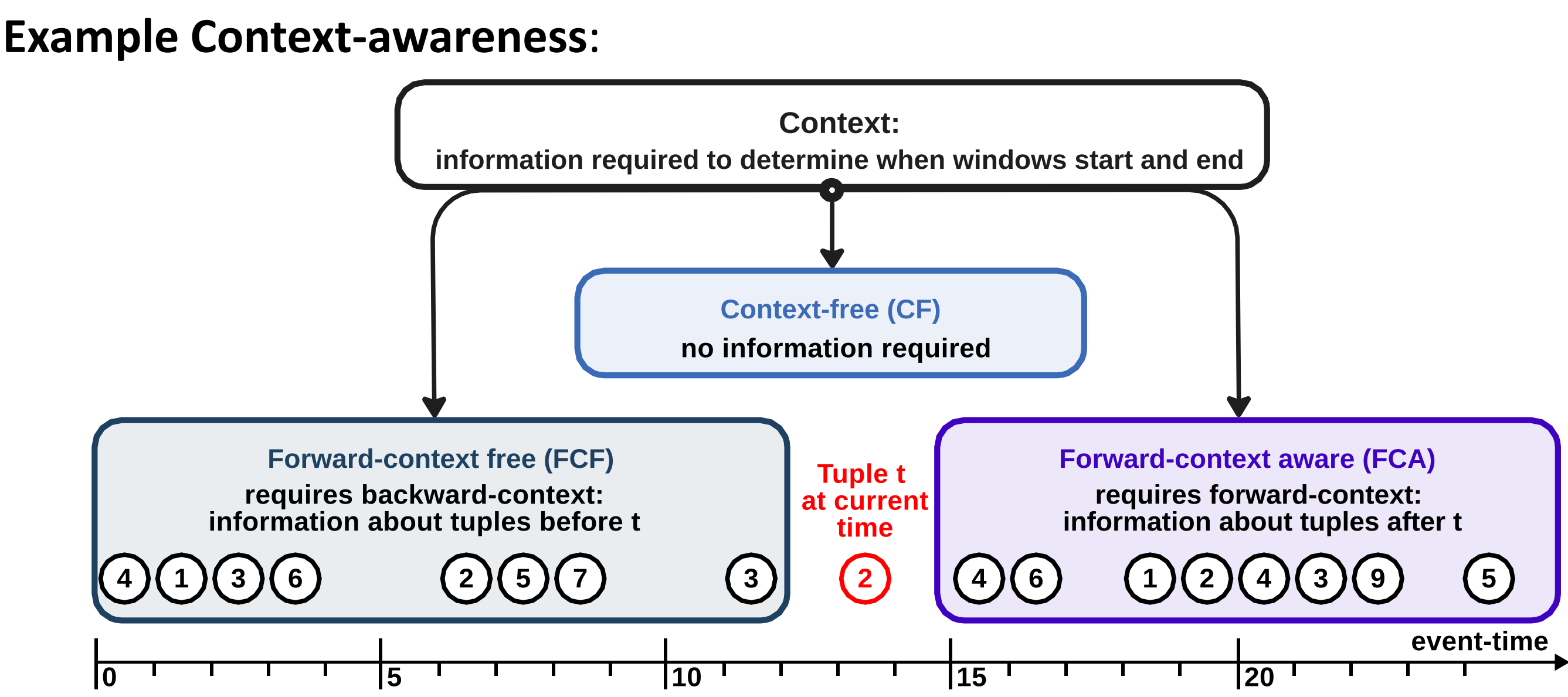
Windowing

Systems divide the unbounded stream into finite subsets of data (i.e., **windows**) to perform aggregations and joins. The window type determines how to split the stream into windows and which tuples are included in the windows. This affects the result of window operations, and enables different application scenarios.



Classification

- We classify window types according to the categories:
- **Window size:** fixed-sized vs. variable-sized
 - **Overlap of individual windows:** overlapping vs. non-overlapping
 - **Window measure:** time-based, count-based, multi-dimensional
 - **Order of inserting and removing tuples:** FIFO vs. non-FIFO
 - **Determinism:**
 - Deterministic: the window operator can immediately decide whether a tuple signals beginning or end of a window
 - Non-deterministic: the window operator requires to process more tuples to make this decision
 - **Context-awareness:** context-free, forward-context free, forward-context aware.



Catalog of Window Types

Window Type	Short Description	Window Size	Overlapping	Window Measure	FIFO	Deterministic	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
Fixed-Band Window	Creates a single window based on pre-defined timestamps for beginning and end	Fixed	No	Time	✓	✓	CF
Landmark Window	Sets one window edge on a specific timestamp (i.e., the landmark)	Fixed	No	Time	✓	✓	CF
Punctuation-based Window	Derives window edges by punctuations embedded in the stream	Variable	No	Time	✓	✓	FCF
Session Window	Ends window after timeout gap when no tuples arrive in a predefined time interval	Variable	No	Time	✓	✓	FCF
Snapshot Window	Starts next window when an interval event starts or ends	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	✓	✗	FCA
Policy-based Window	Provides different policies (configurations) of tumbling and sliding windows	Fixed/Variable	No/Yes	Time/Count/Multi	✓/✗	✓/✗	CF/FCF/FCA
Threshold Frames	Starts next frame when specified attribute is above (or below) a predefined threshold	Variable	No	Time	✓	✗	FCA
Delta Frames	Starts next frame when attribute value changes more than predefined amount	Variable	No	Time	✓	✗	FCA
Boundary Frames	Starts next frame when attribute crosses one of the predefined breakpoints	Variable	No	Time	✓	✗	FCA
Aggregate Frames	Starts next frame when aggregate is greater/smaller than predefined constant	Variable	No	Time	✓	✗	FCA
Adaptive Windowing	Adapts window size dynamically to the rate of change in the data stream	Variable	Yes	Count	✓	✗	FCA
Damped Window	Assigns a weight to each tuple according to an aging function	Fixed	No	Time	✓	✓	FCF
Tilted Window	Stores tuples in summaries at different levels of granularity	Fixed	No	Time	✓	✓	FCF

Systems

	Borealis	STREAM	Telegraph CQ	Naiad	Trill	Amazon Kinesis Data Analytics	Google Cloud Dataflow	IBM Streams	Microsoft Azure Stream Analytics	Microsoft StreamInsight	Apache Beam	Apache Flink	Apache Samza	Apache Spark	Apache Storm	Apache Heron	Scotty
Tumbling Window	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sliding Window	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Fixed-Band Window			✓														✓
Landmark Window			✓									✓					
Punctuation-based Window								✓				✓					✓
Session Window					✓		✓		✓		✓	✓	✓	✓			✓
Snapshot Window										✓							
Slide-by-Tuple Window								✓				✓					✓
Policy-based Window							✓	✓			✓	✓					
Threshold Frames																	✓

Visit our Website
Open-source Window Implementations



<https://scotty.stream/>

