

Window-slicing Techniques Extended to Spanning-event Streams

Aurélie Suzanne - on behalf of co-authors Guillaume Raschia, José Martinez and Damien Tasseti

Polytech Nantes, Université de Nantes - Expandium

23 September 2020


Contact: aurelie.suzanne@ls2n.fr

From points to spanning events

Up to now streaming systems consider only point events

Extension to spanning events:

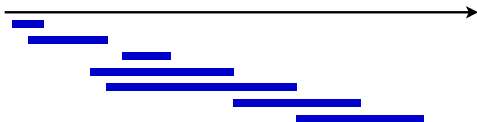
- adapt to real world events
- provide more accurate results

Phone call: 

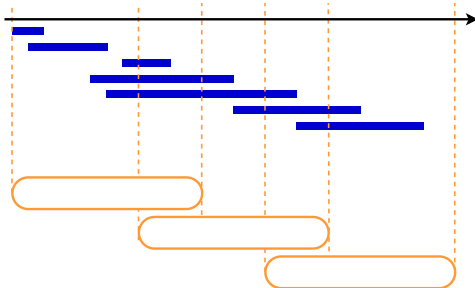
Point event: 

Spanning event: 

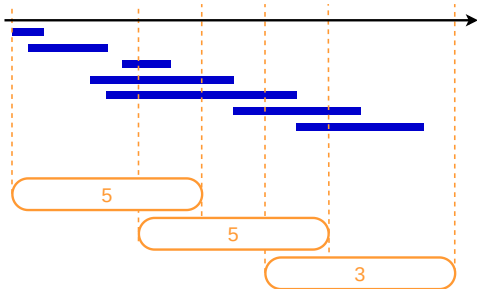
Aggregating data on sliding windows



Aggregating data on sliding windows



Aggregating data on sliding windows

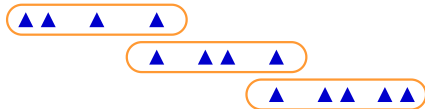


Slicing

Events:



Windows:

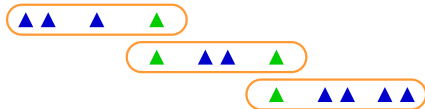


Slicing

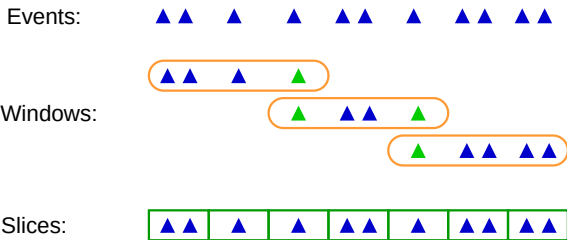
Events:



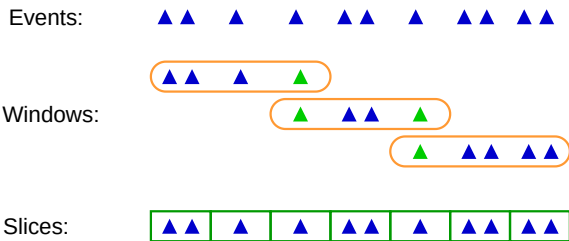
Windows:



Slicing



Slicing

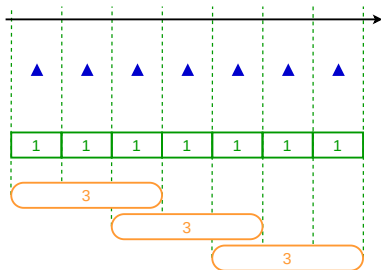


Slicing improves performance:

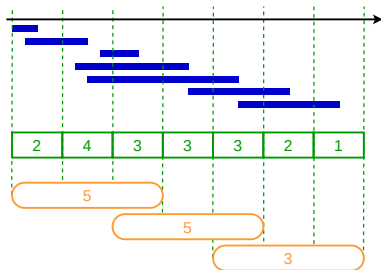
- allows computation sharing among windows
- limits memory usage
- reduces spikes in the system at window release

The problem

INSTANTANEOUS EVENTS:



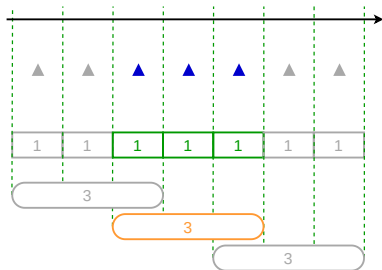
SPANNING EVENTS:



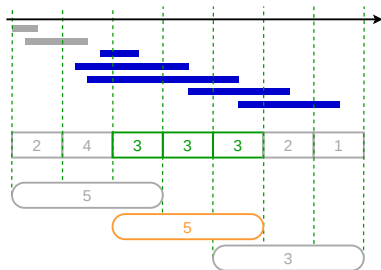
Point event slices cannot be easily reused due to event duplication

The problem

INSTANTANEOUS EVENTS:



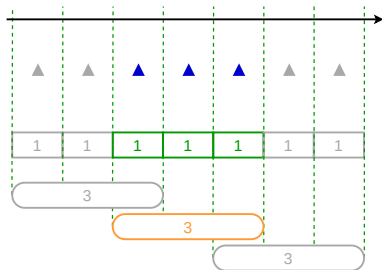
SPANNING EVENTS:



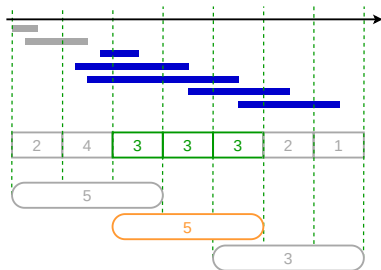
Point event slices cannot be easily reused due to event duplication

The problem

INSTANTANEOUS EVENTS:



SPANNING EVENTS:



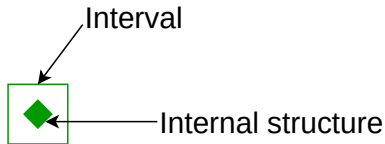
Point event slices cannot be easily reused due to event duplication

Takeaways of this presentation

- Using spanning events in streaming systems
- Adapt slicing techniques to spanning events
- Performance improvements

Point Event Slicing

Slice composition



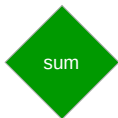
- Slices follow each other without superposition
- Results are kept in pre-computed form

Examples of internal structure

COUNT



SUM



MEAN



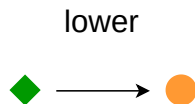
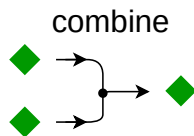
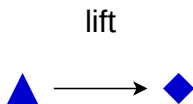
MAX



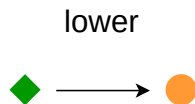
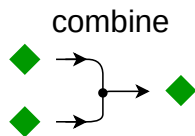
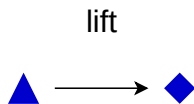
MIN



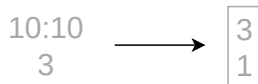
Slicing functions



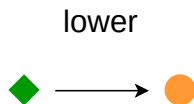
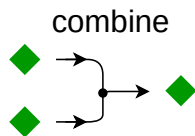
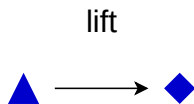
Slicing functions



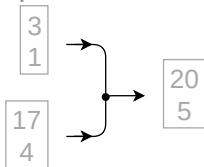
Example: mean lift



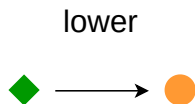
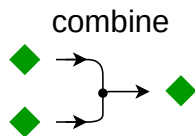
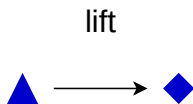
Slicing functions



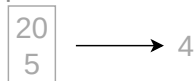
Example: mean combine



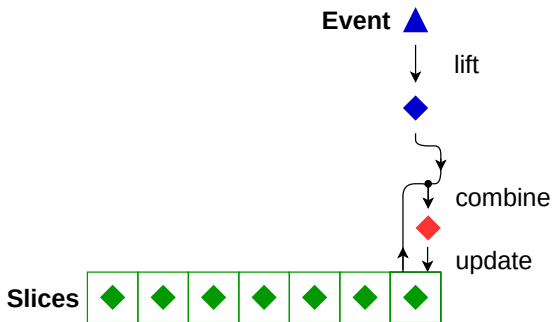
Slicing functions



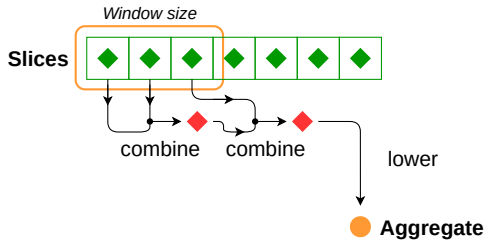
Example: mean lower



Insert event



Release window



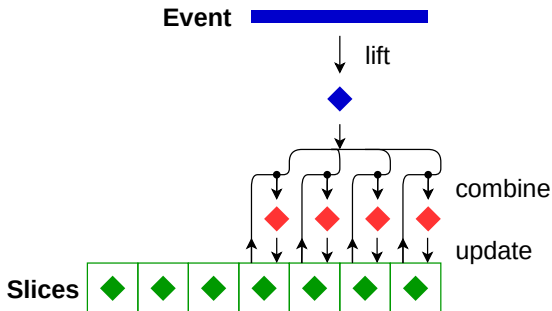
Spanning Events Slicing

Aggregate functions

- **Selective:** max, min, etc.
- **Cumulative:** count, sum, mean, etc.

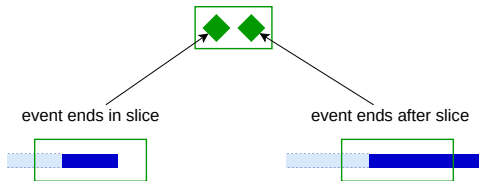
Selective aggregate functions

No duplication problem, we can use classical slices

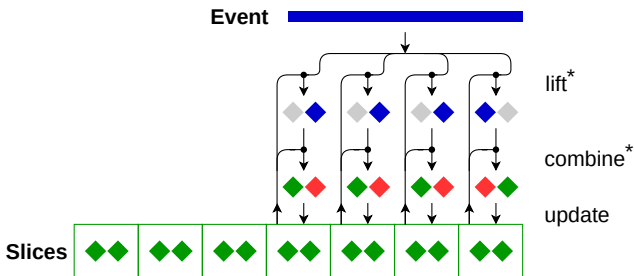


Cumulative aggregate functions

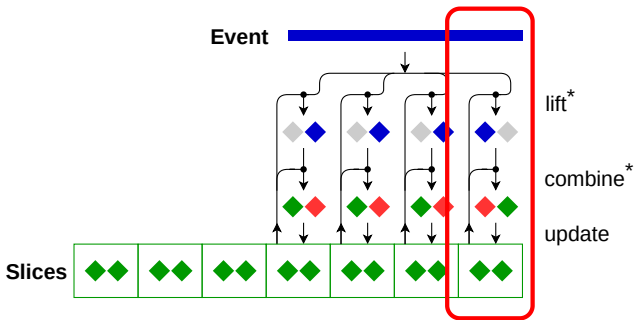
Duplicate internal structure



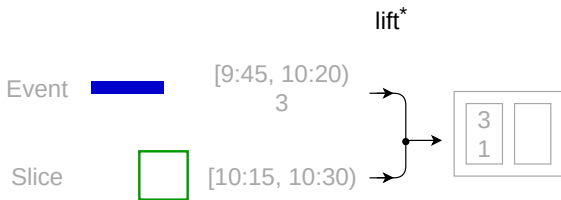
Cumulative aggregate functions: Insert Event



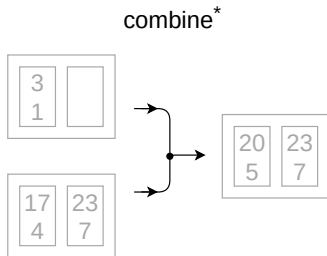
Cumulative aggregate functions: Insert Event



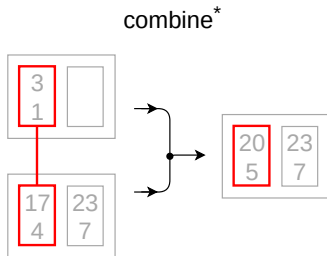
Cumulative aggregate functions: Insert Event



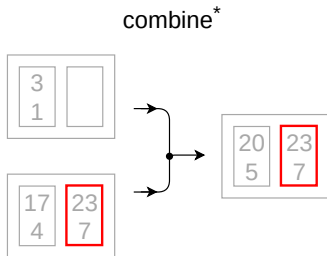
Cumulative aggregate functions: Insert Event



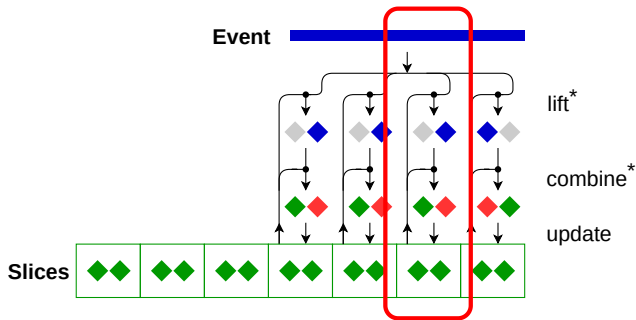
Cumulative aggregate functions: Insert Event



Cumulative aggregate functions: Insert Event



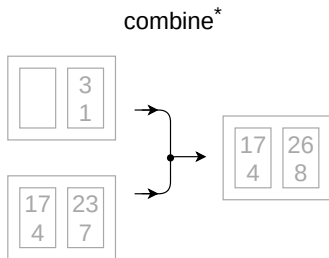
Cumulative aggregate functions: Insert Event



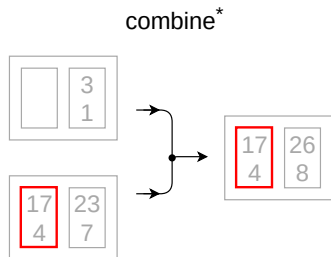
Cumulative aggregate functions: Insert Event



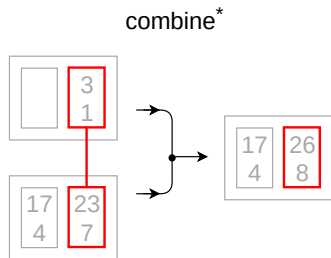
Cumulative aggregate functions: Insert Event



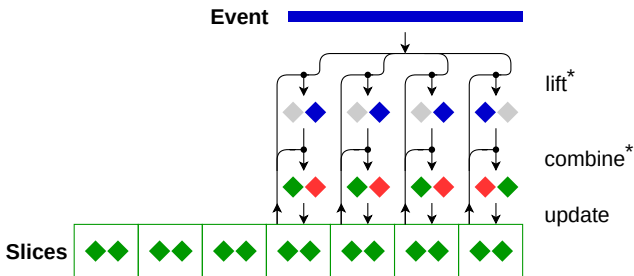
Cumulative aggregate functions: Insert Event



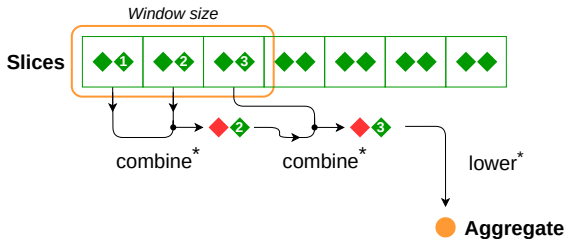
Cumulative aggregate functions: Insert Event



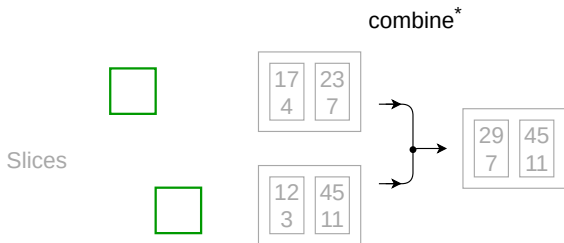
Cumulative aggregate functions: Insert Event



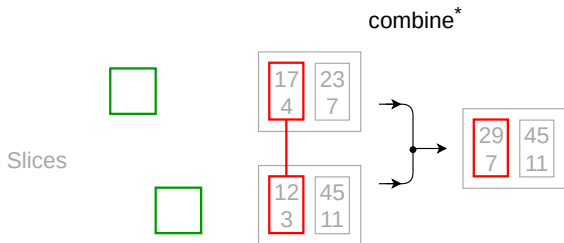
Cumulative aggregate functions: Release Window



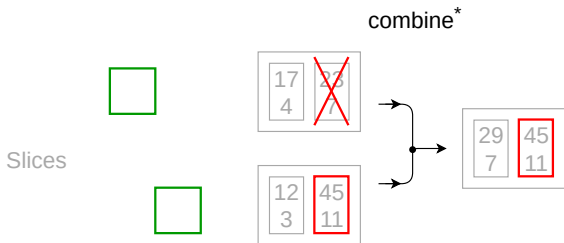
Cumulative aggregate functions: Release Window



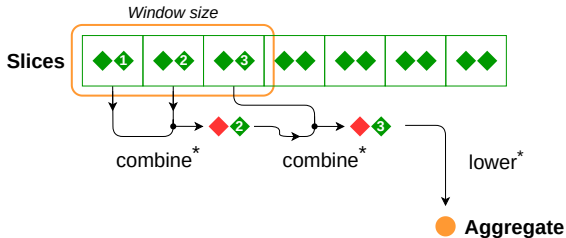
Cumulative aggregate functions: Release Window



Cumulative aggregate functions: Release Window

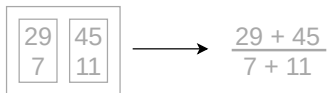


Cumulative aggregate functions: Release Window



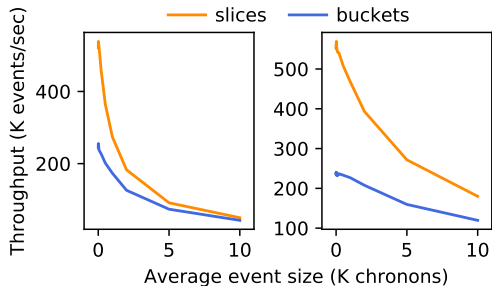
Cumulative aggregate functions: Release Window

lower*



Experiments

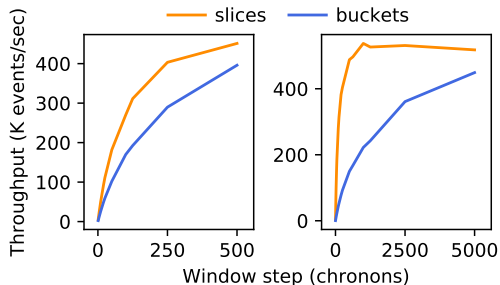
Event Size



(a) Window range ω and step β :

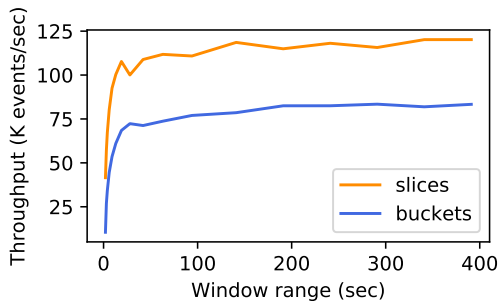
$(\omega = 1K, \beta = 0.2K)$ vs. $(\omega = 10K, \beta = 2K)$

Overlap



(b) Window range $\omega = 500$ vs. $\omega = 5000$

Telecommunication dataset



(c) Window step $\beta = \omega/5$

Conclusion

Conclusion

Streaming systems can be extended to use spanning events

Slicing technique can be used with spanning events:

- Spanning events duration induces duplication
- Specific structures to moderate this constraint
- Better performance than with naive technique

Conclusion

Streaming systems can be extended to use spanning events

Slicing technique can be used with spanning events:

- Spanning events duration induces duplication
- Specific structures to moderate this constraint
- Better performance than with naive technique