Agenda

- Introduction on Stream Processing Models [done]
- Declarative Language: Opportunities, and Design Principles
- Comparison of Prominent Streaming SQL Dialects for Big Stream Processing Systems
- Conclusion

"The key idea of declarative programming is that a **program** is a **theory** in some suitable **logic**, and the **computation** is **deduction** from the **theory**"

–J.W. Lloyd

Advantages

- Decouple interpretation and execution (e.g. parallelism)
- Allows optimisation relying on the formal semantics
- IDEALLY PORTABLE (well-defined semantics)

How to design a good language?

Minimality

a language should provide only a small set of needed language constructs so that the same meaning cannot be expressed by different language constructs;



Symmetry

a language should ensure that the same language construct always expresses the same semantics regardless of the context



Orthogonality

a language should guarantee that every meaningful combination its constructs is applicable.



When do we need it?

- Writing the optimal solution is as hard as solving the problem (e.g. JOIN optimisation)
- We want to enhance programmer productivity by adding Domain-Specific abstraction (e.g. streams)
- We want to limit the expressiveness of the languages to ensure some nice property (e.g. decidability)







SELECT name
FROM (
SELECT id, name
FROM People) p
WHERE p.id = 1

Parsing

- Obtaining the Declarative Program/Query
- Verify it is is syntactically valid
- Creating an AST







Logical Planning

- Obtaining the AST of the program/query
- Verify all the preconditions hold
- Apply optimisations
- Errors: statistics not updated, wrong decision
- Generates logical plan





Example





Physical Planning

- Obtaining the logical plan of the program/query
- Verify all the preconditions
- Errors: table not exists
- Generates physical plan





results

Example





Executing

- Obtain physical plan of the query
- Load it for execution
- Run!



Runtime Errors

- Input not compliant to the expected one
- Table dropped while running
- Network fail (fixable)
- Node fail (fixable)

