Centre for
Data Analytics

Insight

IoT-Intelligence and Web Stream Reasoning

Towards Complex Reasoning over Big Data Streams with Answer Set Programming

Alessandra Mileo

Senior Research Fellow

Insight Centre for Data Analytics, NUIG

alessandra.mileo@insight-centre-org

Ali Intizar

Research Fellow

Insight Centre for Data Analytics, NUIG

ali.intizar@insight-centre-org

CityPulse

IoT- Intelligence at *Unit for Reasoning & Querying (URQ)*

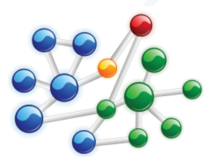
1. Representation and linking 
2. Finding what we need 
3. Dynamic Problem Solving (Stream Reasoning) 

Centre for
Data Analytics

CityPulse

Insight

IoT- Intelligence at *Unit for Reasoning & Querying (URQ)*

1. Representation and linking 
2. Finding what we need 
3. Dynamic Problem Solving (Stream Reasoning) 

Smart Cities

Enterprise Communication

What streams do I need and how "good" are they? (Stream Discovery and Federation)

Quality and context-aware stream discovery

- What information do I need?
  – Data interoperability: Semantic descriptions
  – Interface interoperability: streams as event services
- How good is it?
  – ADAPT to quality requirements and preferences for data source selection
  – Efficient processing of event logic

Quality and context-aware stream discovery

Semantic Web

- What information do I need?
  - Data interoperability: Semantic descriptions
  - Interface interoperability: streams as event services
- How good is it?
  - ADAPT to quality requirements and preferences for data source selection
  - Efficient processing of event logic

Quality and context-aware stream discovery

W3C WORLD WIDE WEB consortium

Semantic Web

Service Oriented Architectures

- What information do I need?
  - Data interoperability: Semantic descriptions
  - Interface interoperability: streams as event services
- How good is it?
  - ADAPT to quality requirements and preferences for data source selection
  - Efficient processing of event logic

Quality and context-aware stream discovery

- What information do I need?
  - Data interoperability: Semantic descriptions
  - Interface interoperability: streams as event services
- How good is it?
  - ADAPT to quality requirements and preferences for data source selection
  - Efficient processing of event logic

W3C WORLD WIDE WEB consortium

Semantic Web

Service Oriented Architectures

Continuous constraint checking

Quality and context-aware stream discovery

- What information do I need?
  - Data interoperability: Semantic descriptions
  - Interface interoperability: streams as event services
- How good is it?
  - ADAPT to quality requirements and preferences for data source selection
  - Efficient processing of event logic

Semantic Web

Service Oriented Architectures

Continuous constraint checking

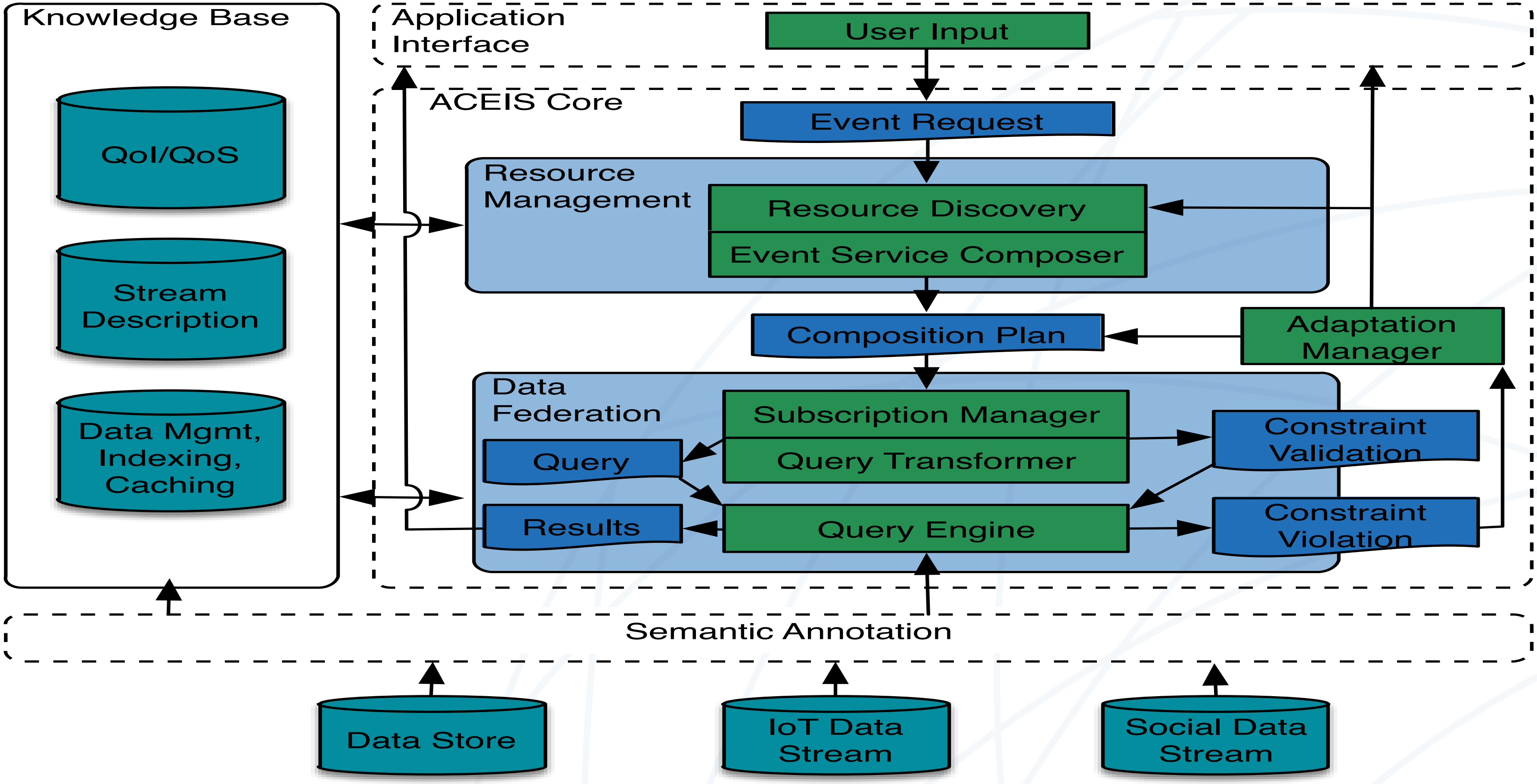
DSMS and CEP

Summary of the Approach

- How to describe complex event services?
  - Create an Event Service Ontology with Event Patterns.
- How to determine if two event patterns are functionally equivalent?
  - Create and compare canonical event patterns to find substitutes.
- How to create event compositions and choose the optimal?
  - Top-down traverse to find functionally-equivalent canonical patterns.
- How to derive event service compositions efficiently?
  - Construct and utilize an Event Reusability Hierarchy for event service composition.

Gao, F., Ali, M.I., Curry, E., Mileo, A.: *On Discovery and Integration of Urban Data Streams for Realtime Smart City Applications*. J. Data Semantics (2015) to appear

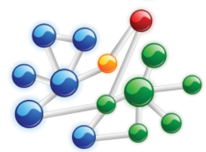
Automated Complex Event Implementation System



IoT- Intelligence at *Unit for Reasoning & Querying (URQ)*

1. Representation and linking

2. Finding what we need

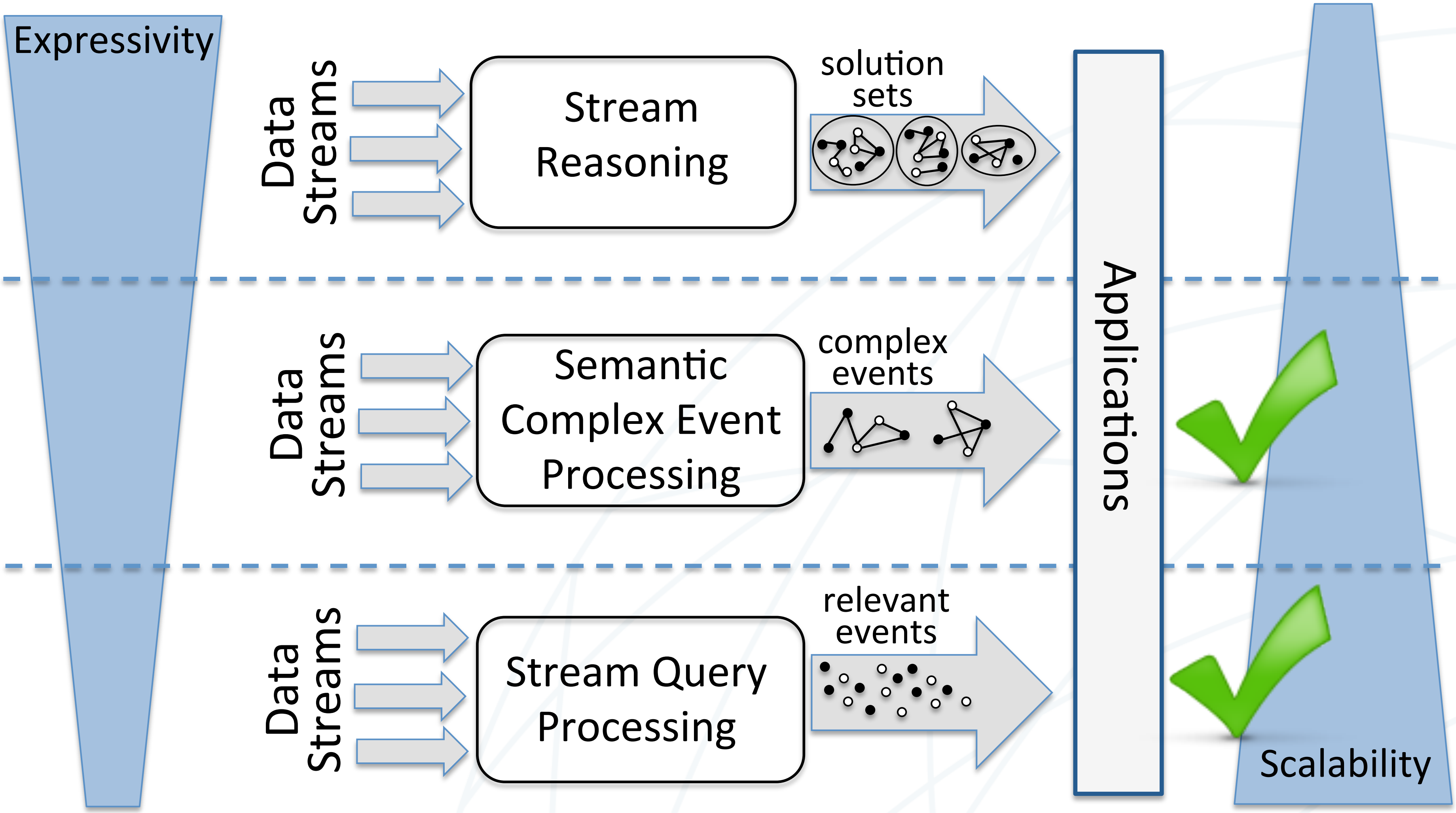


3. Dynamic Problem Solving (Stream Reasoning)

Smart Cities

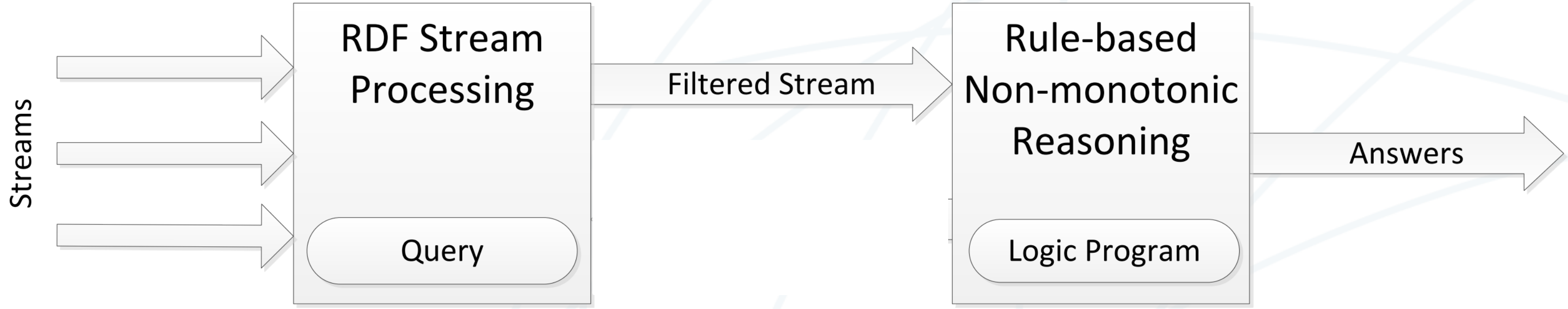
Enterprise Communication

How to leverage the IoT and Semantic Web infrastructure for (efficient) Web Stream Reasoning?


Expressivity
Data Streams
Stream Reasoning
solution sets
Data Streams
Semantic Complex Event Processing
complex events
Data Streams
Stream Query Processing
relevant events
Applications
Scalability

The StreamRule idea

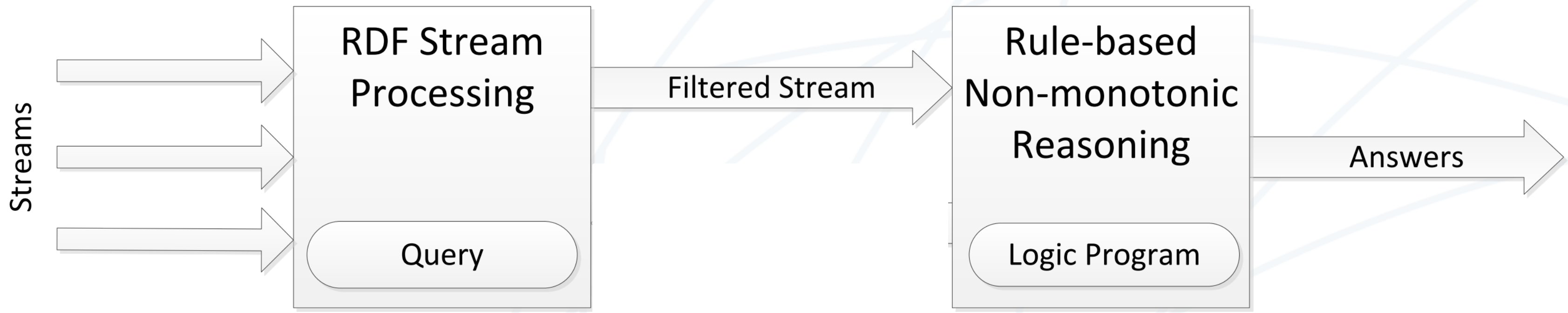
- 2-tier approach: not all data streams are relevant for complex reasoning
- Enrich the ability of complex reasoning over data streams
- Keep the solution scalable
- Leverage existing engines from both stream processing and non-monotonic reasoning research areas



Mileo, A., Abdelrahman, A., Policarpio, S., Hauswirth, M.: Streamrule*: A nonmonotonic stream reasoning system for the Semantic Web.* In: RR 2013, 247–252

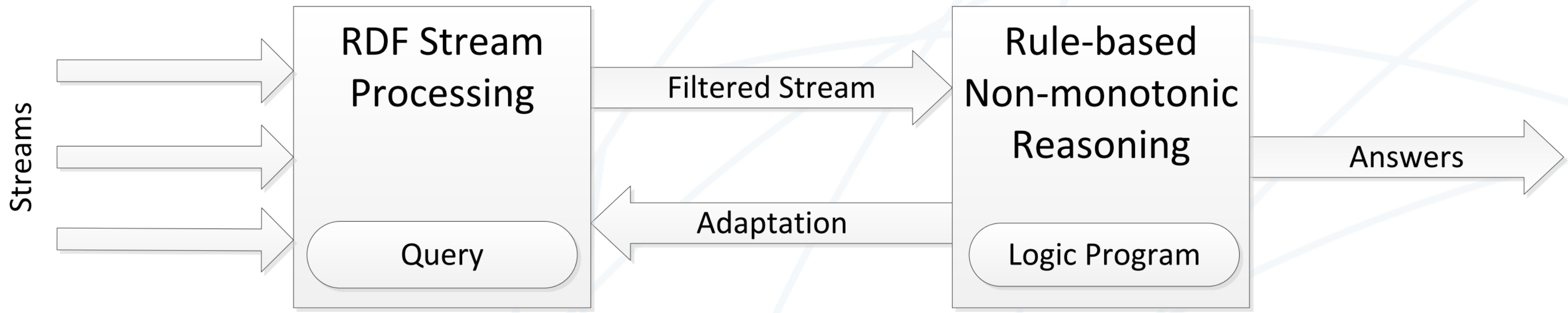
Limitations

- The more expressive the inference task, the longer it takes to perform reasoning
- Bottleneck when results are returned not as fast as the next input arrives



Limitations

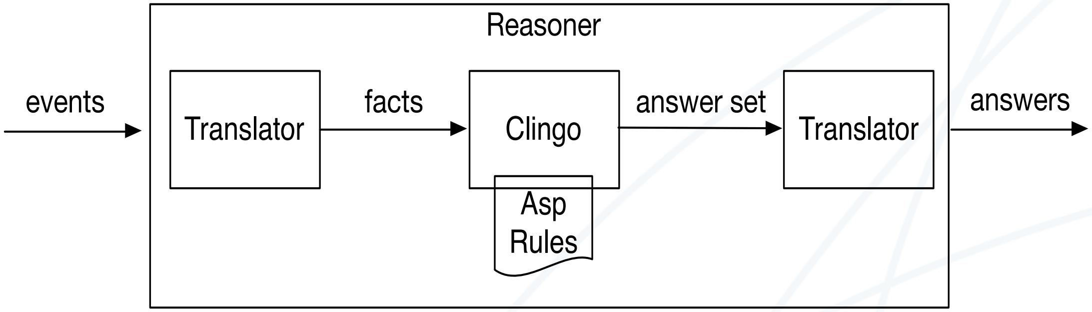
- The more expressive the inference task, the longer it takes to perform reasoning
- Bottleneck when results are returned not as fast as the next input arrives



Adaptation Heuristics: ongoing work

- More than an engineering problem
  – How to model interactions between RSP and ASP components, including different semantics, input split, window-size tuning,…
- Design and runtime features
  – E.g. operational semantics (design) and throughput (runtime)
- Streaming rate and window size: where's the tradeoff?
- Reasoning Complexity: how far can we go? How can we parametrise the complexity to estimate the execution time

Goal



Given a fixed streaming size S with fixed complexity C and unit of time U, find a window size W such that the time required to process S events using windows of size W is less than or equal to one unit of time

$$T_{\omega}(S,W) \leq U$$

Germano, S., Pham, T.L., Mileo, A.: *Web stream reasoning in practice: on the expressivity vs. scalability tradeoff.* In: Web Reasoning and Rule Systems - 9th International Conference, RR 2015, Berlin, Germany, August 4-5, 2015, Proceedings. (2015) 105–112

Experimental setup

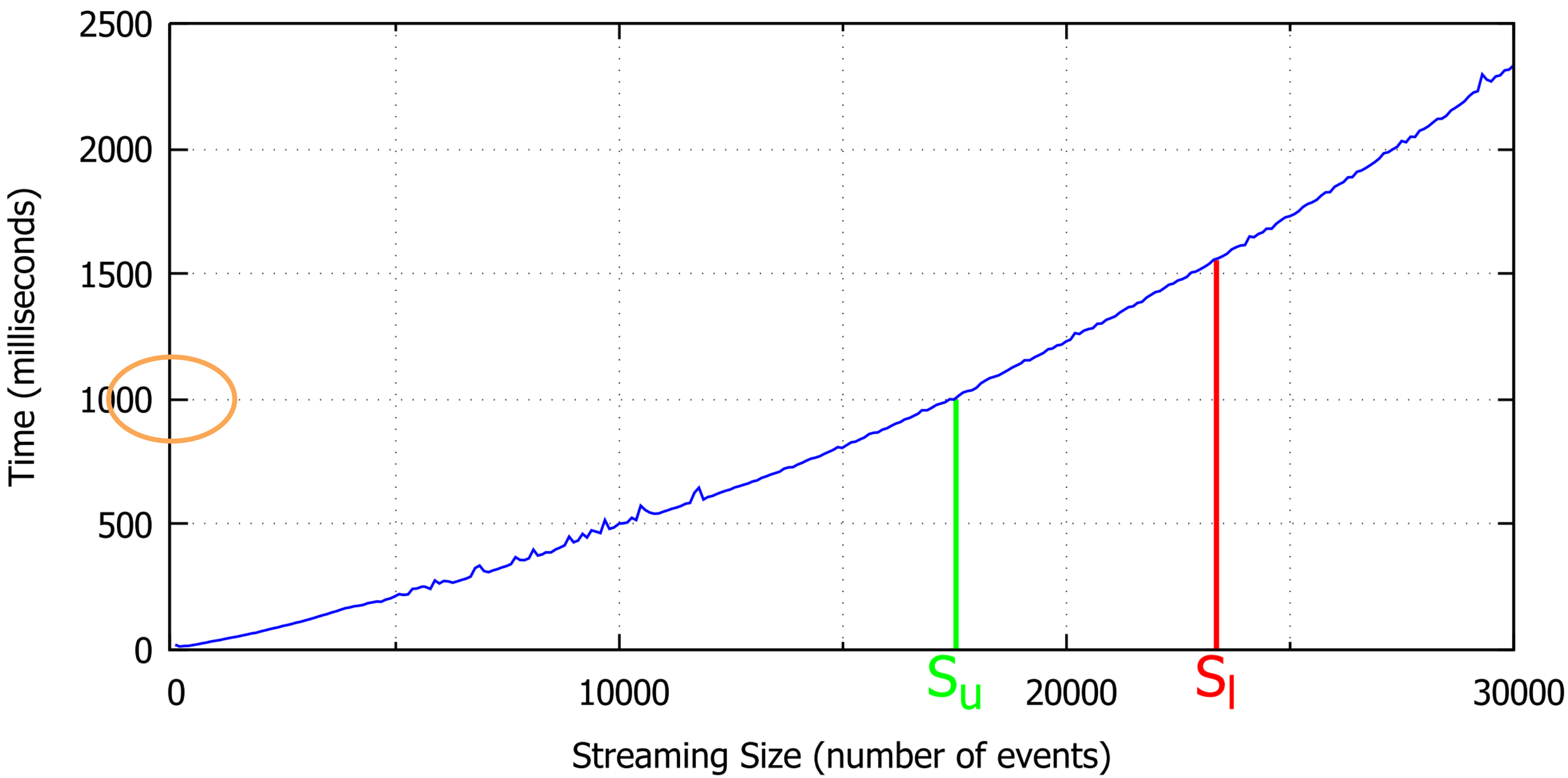
- Dataset
  - Simulated randomly generated events of the type
    event(type, name, value, latitude, longitude)
    E.g. event(weather, strong-wind, 2014-11-26T13:00:00, 38.011736, 12.186724)
- Reasoning tasks
  - Ranking event criticality
  - Contextualizing events based on user status
  - Default rule to detect changes in event criticality
- Run
  - Streaming size up to 30000
  - Reasoner triggered 20 times for each S

Centre for
Data Analytics

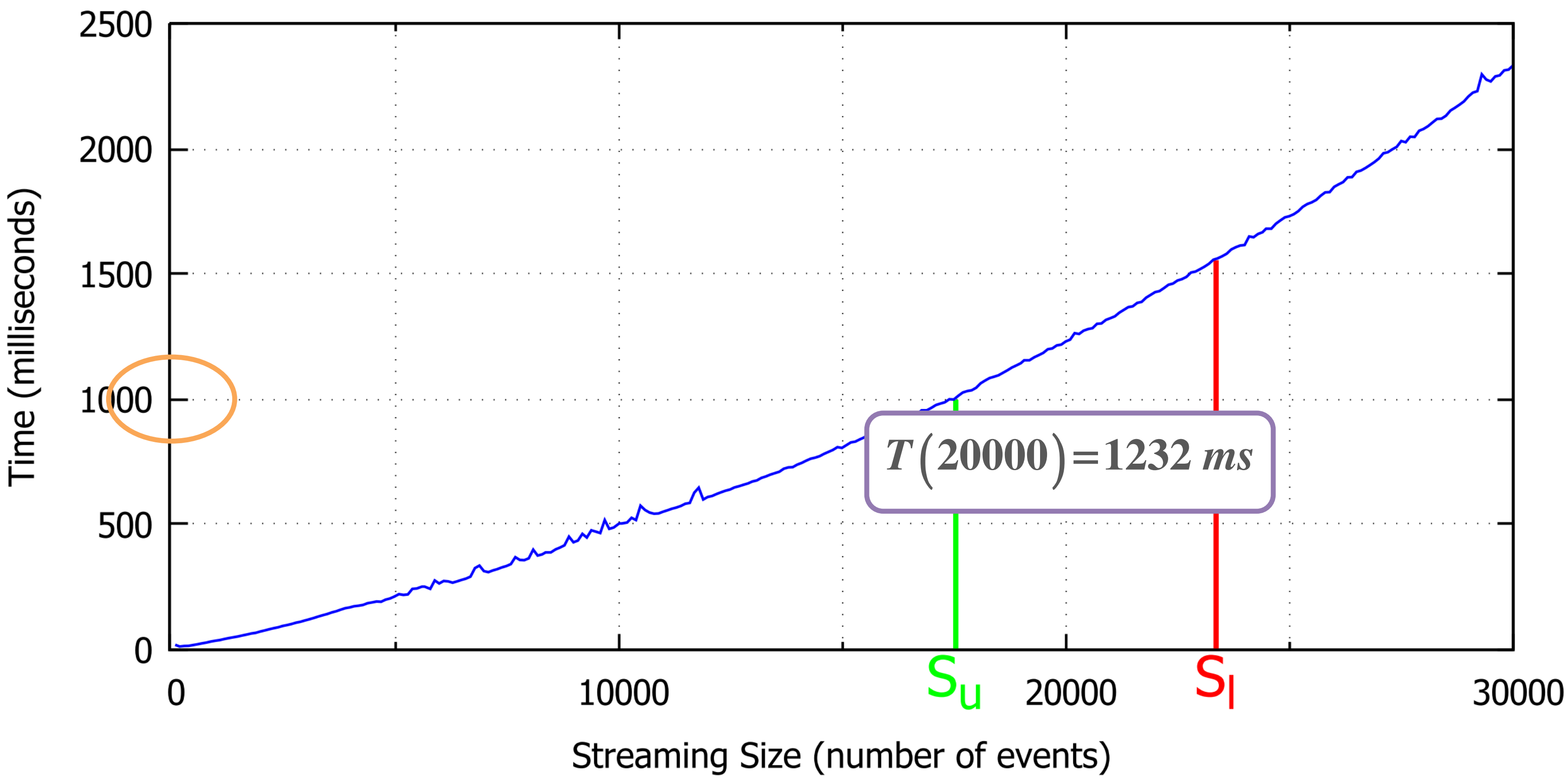
CityPulse

Insight

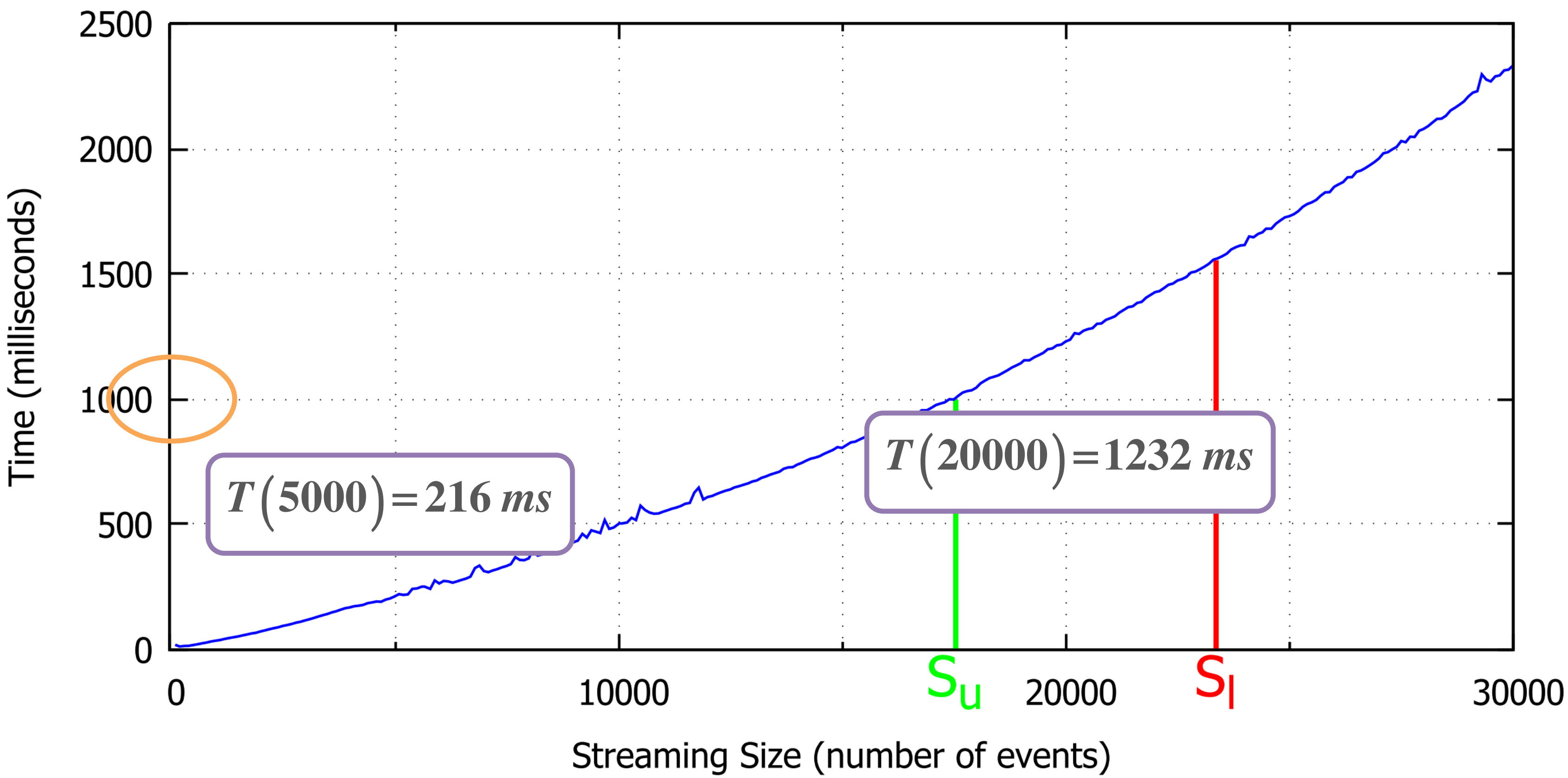
Empirical Results



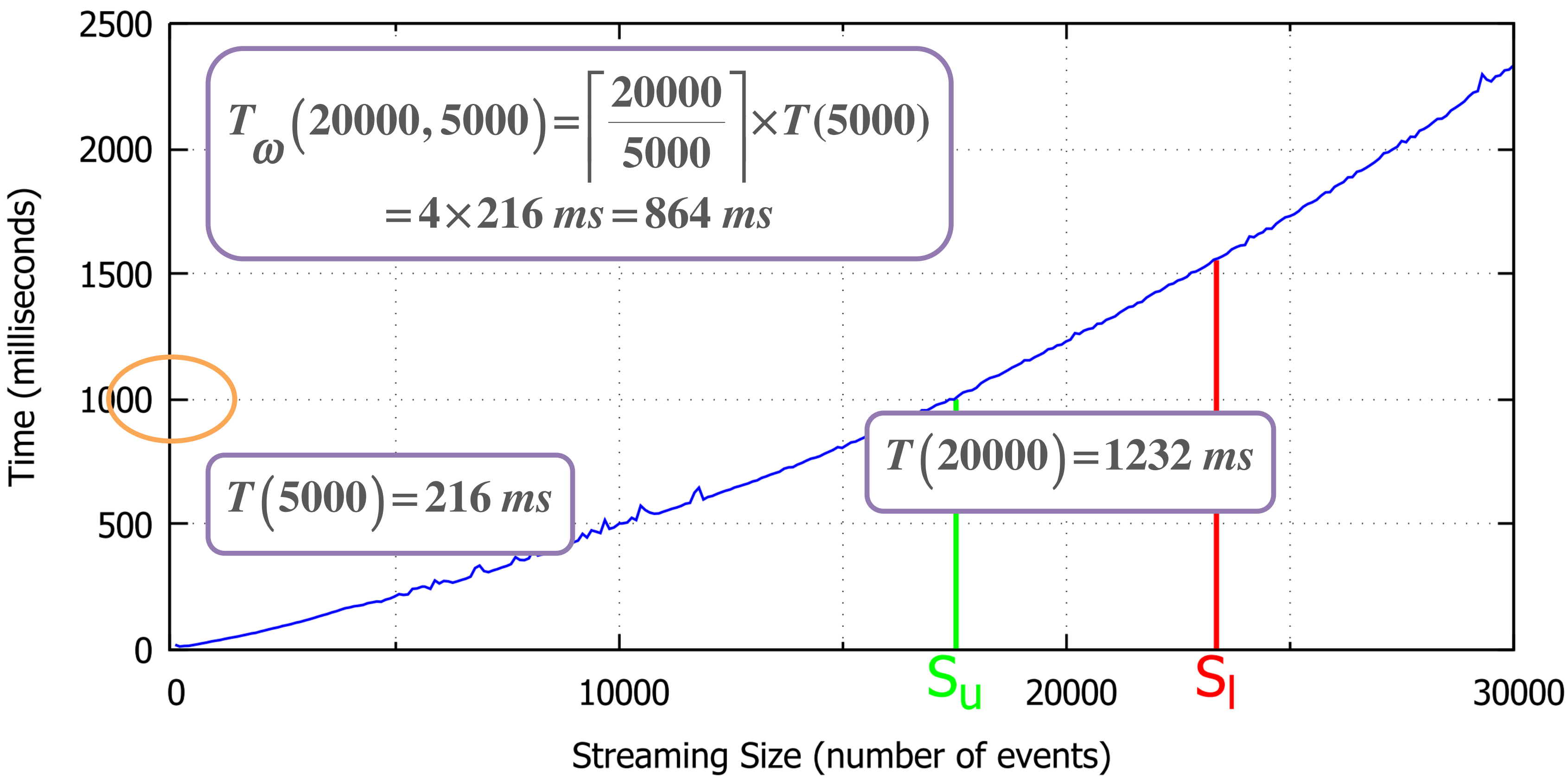
Empirical Results



Empirical Results

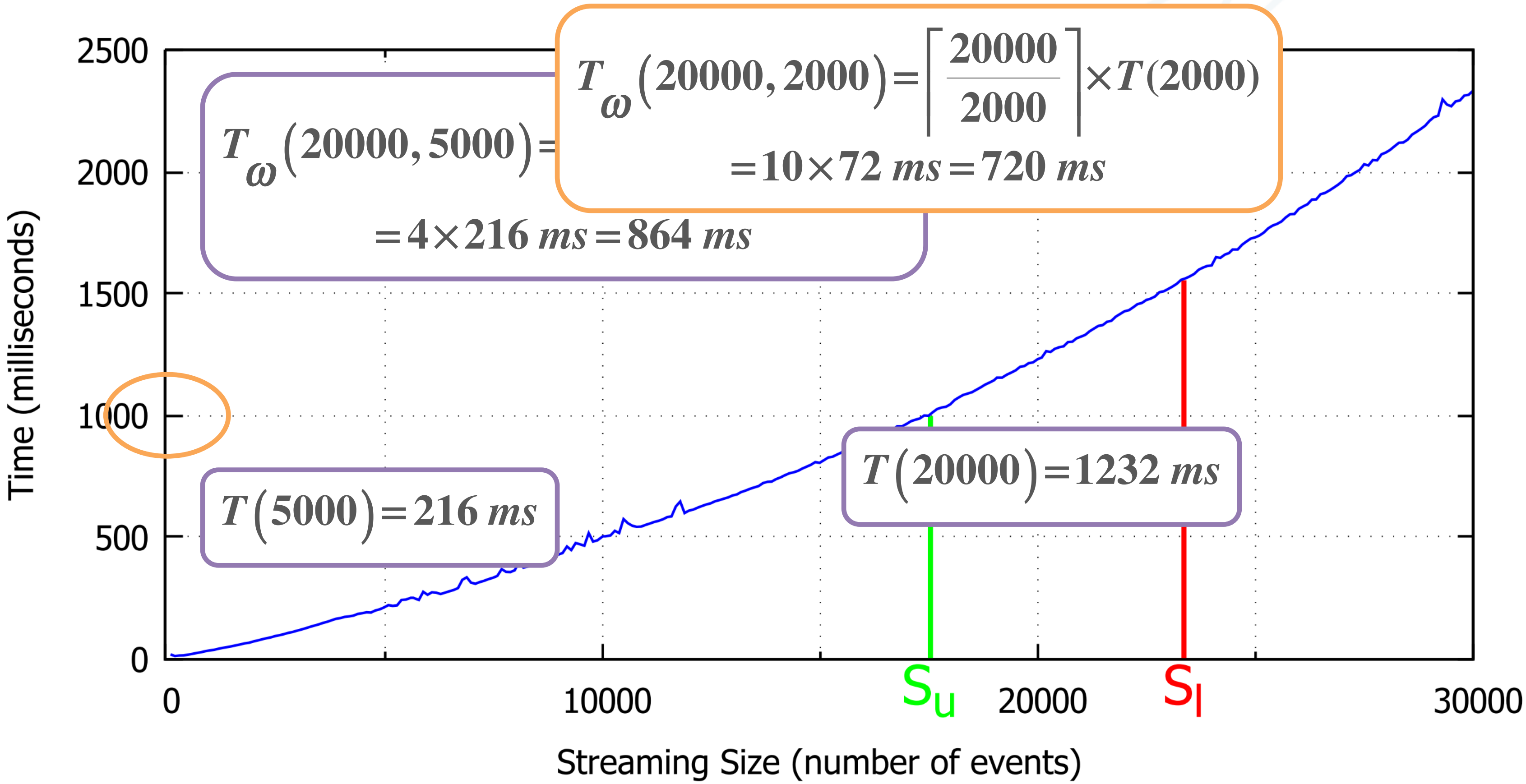


Empirical Results



Centre for
Data Analytics

Empirical Results



Ongoing work

- Relaxing the independence assumption
  - Extended notion of dependency graph
  - Possibly use duplication
  - Smaller input sets given to the same ASP program
  - Demonstrate correctness of results
- Going parallel
  - Explore parallelism of SPARK for higher scalability
  - Requires to map an ASP program to SPARK jobs
- Correlation between reasoning complexity and execution time

Dealing with Uncertainty and learning relational structures

IoT data are messy: deal with uncertainty

- Expressive inference
  - non-monotonicity, noisy, partial and inconsistent data
- "ease of" declarative logic-based reasoning to model a problem/ domain. Still we need to manage uncertainty and non-monotonicity
- Probabilistic rules for uncertain knowledge and learning by example
  - represent, use, infer and learn probabilistic knowledge (PrASP)

Can we (*learn the*) answer to questions about *uncertain knowledge* using *qualitative (declarative) inference* in *dynamic environments*?

What is Streaming PrASP

A framework that uses:

1. PrASP as an uncertainty reasoning server to reason over Streaming Web Data
2. Continuous Query Processing over Linked Data Streams for data filtering

Nickles, M., Mileo, A.: *Web stream reasoning using probabilistic answer set programming.* In: Web Reasoning and Rule Systems (RR) 2014. (2014) 197–205

What is Streaming PrASP

A framework that uses:

1. PrASP as an uncertainty reasoning server to reason over Streaming Web Data
2. Continuous Query Processing over Linked Data Streams for data filtering

Nickles, M., Mileo, A.: *Web stream reasoning using probabilistic answer set programming.* In: Web Reasoning and Rule Systems (RR) 2014. (2014) 197–205

What is PrASP then?

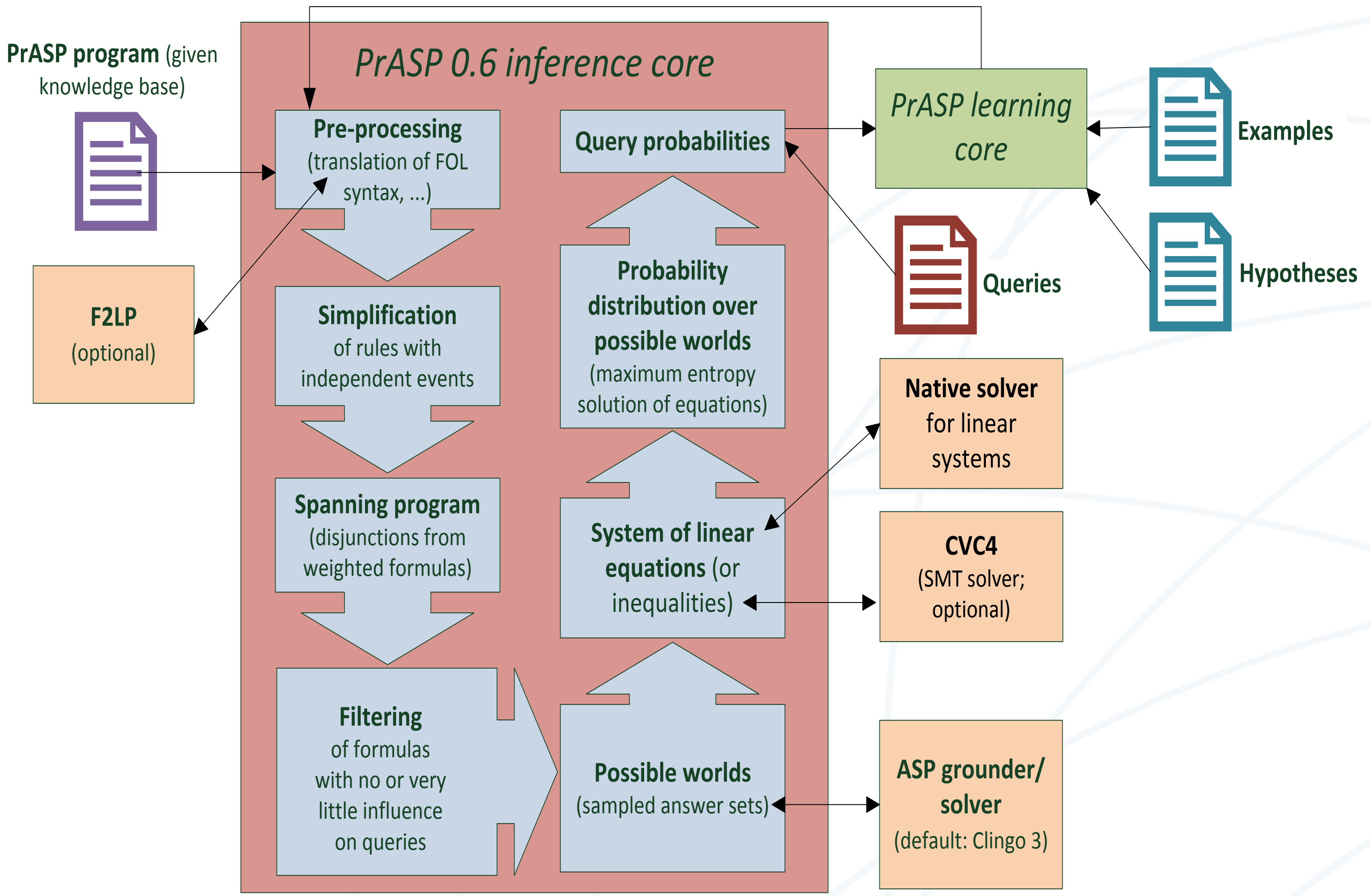
PrASP is…

… an experimental Statistical Relational Learning (SRL) reasoner based on *Answer Set Programming* (ASP)

PrASP can…

… represent, use infer and learn probabilistic knowledge

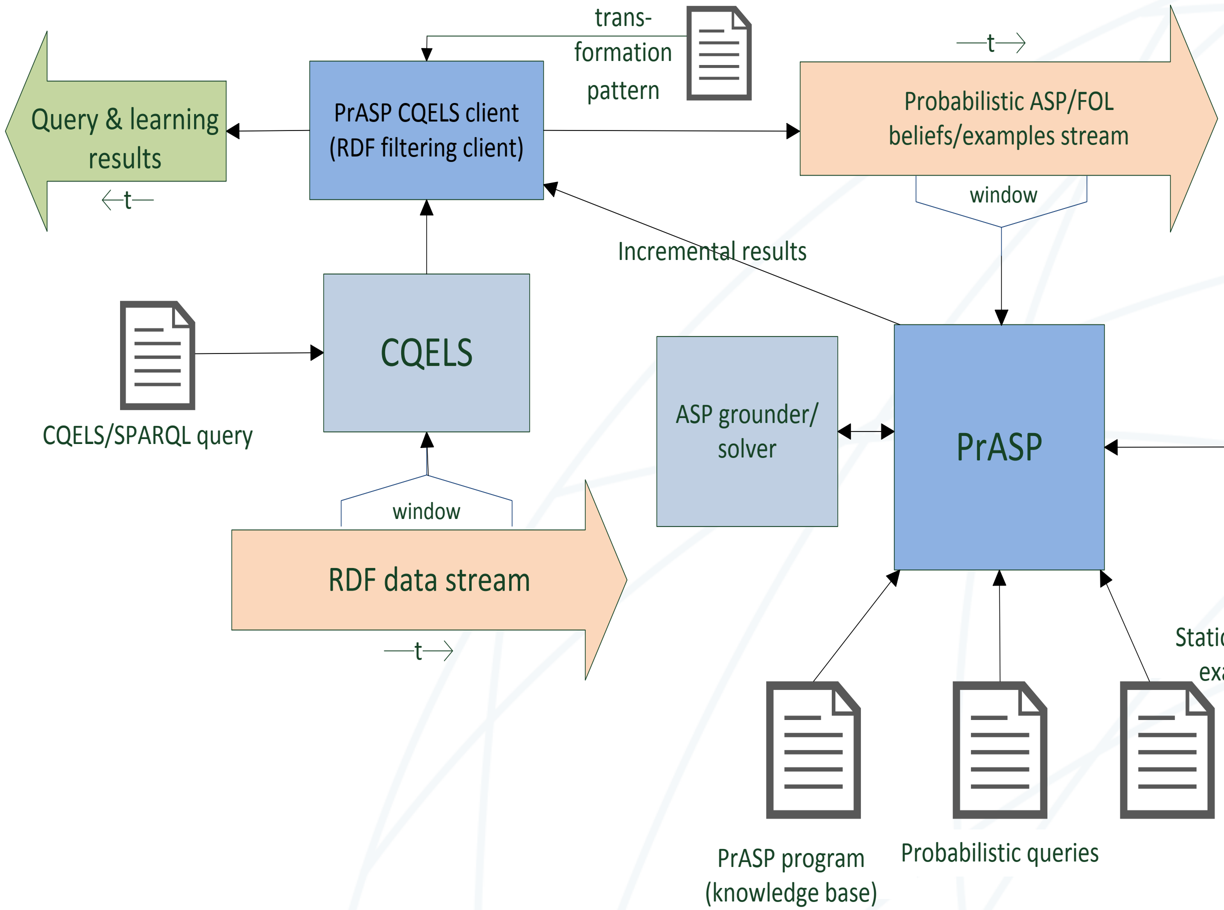
Matthias Nickles, Alessandra Mileo: *A System for Probabilistic Inductive Answer Set Programming.* SUM 2015: 99-105


PrASP program (given knowledge base)
F2LP
(optional)
PrASP 0.6 inference core
Pre-processing
(translation of FOL syntax, ...)
Simplification
of rules with independent events
Spanning program
(disjunctions from weighted formulas)
Filtering
of formulas with no or very little influence on queries
Query probabilities
Probability distribution over possible worlds
(maximum entropy solution of equations)
System of linear equations (or inequalities)
Possible worlds
(sampled answer sets)
PrASP learning core
Examples
Hypotheses
Queries
Native solver
for linear systems
CVC4
(SMT solver; optional)
ASP grounder/ solver
(default: Clingo 3)

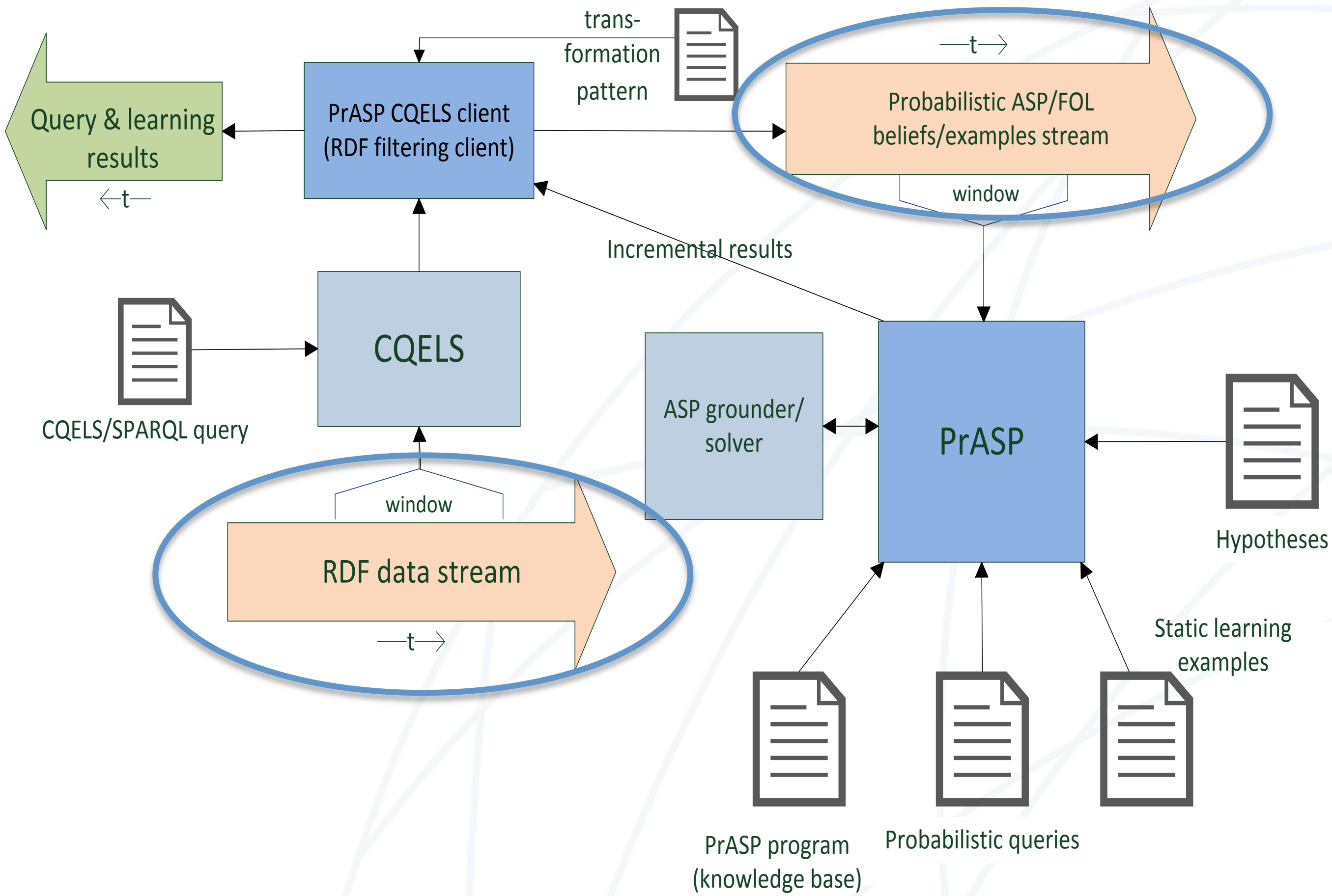
Streaming PrASP

- Streaming ***new beliefs*** are added incrementally to a loaded PrASP program
- Streaming ***new learning examples*** are added to the set of learning examples ***E***
- Assert/Retract, time decay and sliding windows supported
- Windows prefixes realized by a caching mechanism (no reactive ASP used) for faster processing
- Preprocessing based on RDF query processing over streams (SPARQL 1.1 + streaming operators)

Streaming PrASP framework

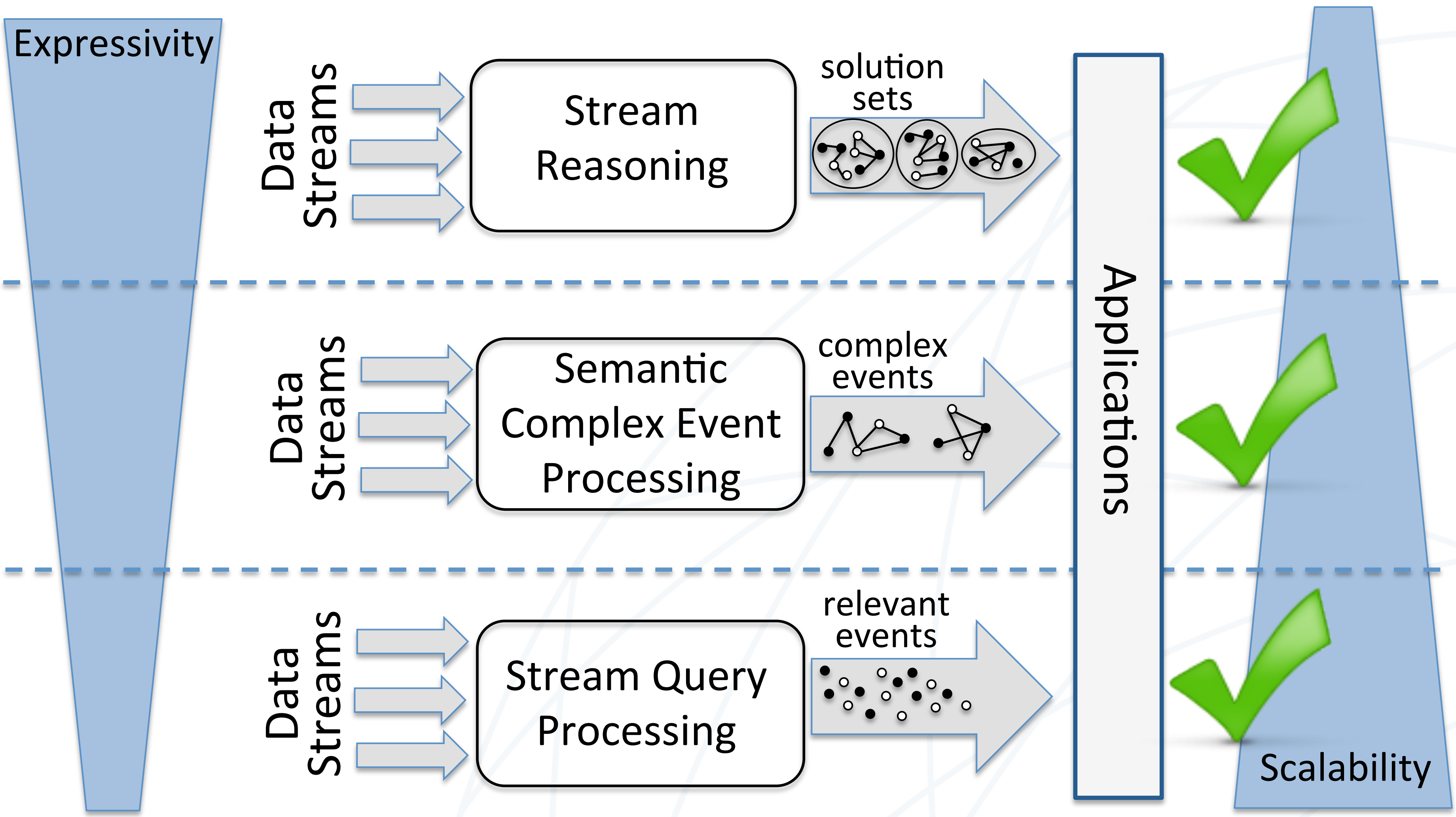


Streaming PrASP framework



Ongoing work

- More experiments on Web Stream Reasoning with PrASP
  - Looking for Usecase Scenarios and DATA to test PrASP
  - Building a set of modules that can be downloaded and used for feedback
- Continuously exploring options for optimization, especially in the learning task
- More ambitious goals
  - Structure Learning
  - Relation between Streaming ASP and PrASP (currently not using Streaming ASP)


Expressivity
Data Streams
Stream Reasoning
solution sets
Data Streams
Semantic Complex Event Processing
complex events
Data Streams
Stream Query Processing
relevant events
Applications
Scalability

Centre for
Data Analytics

Insight

IoT-Intelligence and Web Stream Reasoning

Towards Complex Reasoning over Big Data Streams with Answer Set Programming

Alessandra Mileo
Senior Research Fellow
Insight Centre for Data Analytics, NUIG
alessandra.mileo@insight-centre-org

Ali Intizar
Research Fellow
Insight Centre for Data Analytics, NUIG
ali.intizar@insight-centre-org

CityPulse

DCU

NUI Galway
OÉ Gaillimh

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

UCD
DUBLIN

sfi