
Storing and Querying the Valid Time of Triples in Linked Geospatial Data

Konstantina Bereta, Panayiotis Smeros, Manolis Koubarakis

Konstantina.Bereta@di.uoa.gr

ESWC 2013, Montpellier



Dept. of Informatics and Telecommunications,
National and Kapodistrian University of Athens, Greece

Outline

- **Introduction**
- The “t” of stRDF and stSPARQL
- Implementation in Strabon
- Experimental evaluation
- Future Work and Conclusions

Time dimensions in linked geospatial data

- **User-defined time:** A time value (literal) with no special semantics.
- **Valid time:** The time when a fact (represented by a triple) is true in the modeled reality.
- **Transaction time:** The time when the triple is current in the database.

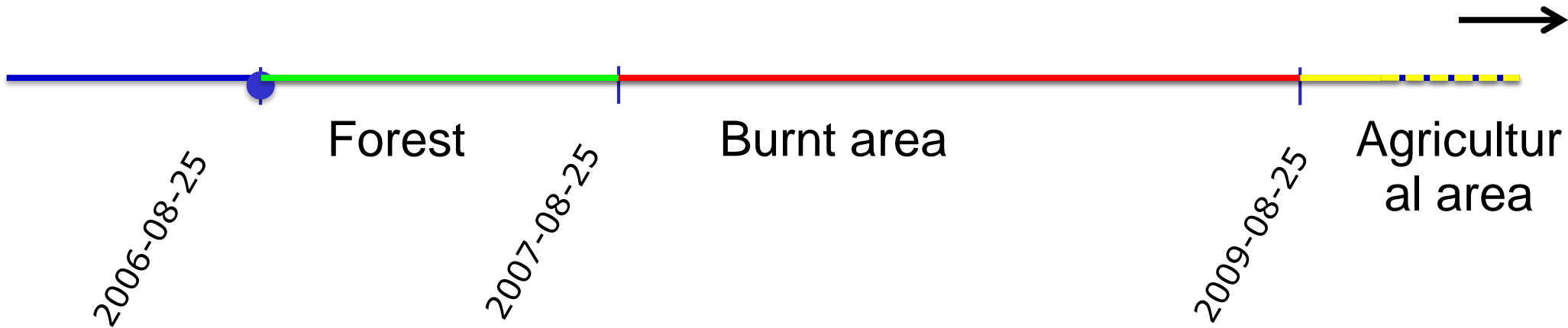
Example of valid time



Forest

Burnt area

Agricultural
area



Previous work

- **Time in relational databases**
 - TSQL2 (1995)
 - SQL:2011 (Oracle Workspace manager, IBM DB2, Teradata)
- **Time in RDF**
 - Temporal RDF (Gutierrez et al., ESWC 2005)
 - Annotated RDFS (Lopes et al., ISWC 2010)
 - Mapping to standard RDF:
 - Reification
 - Named graphs (Tappolet and Bernstein, ESWC 2009)

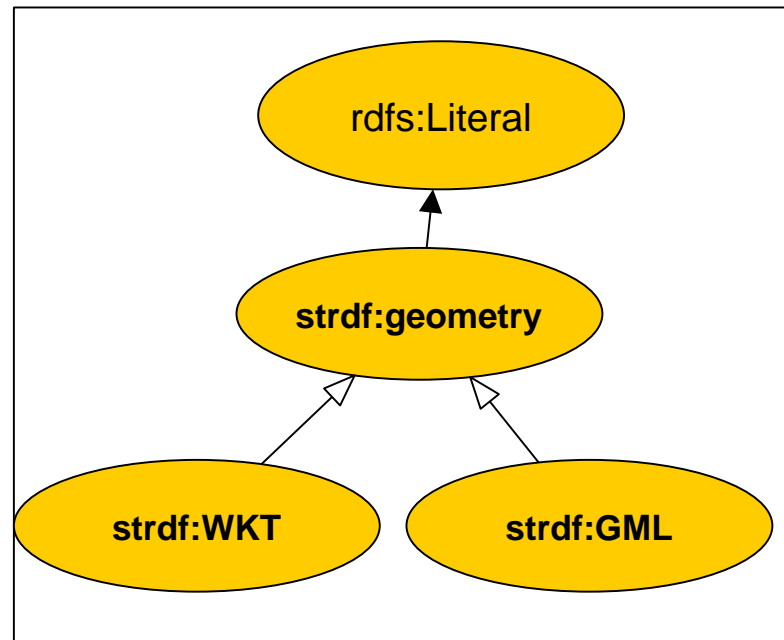
Outline

- Introduction
- **The “t” of stRDF and stSPARQL**
- Implementation in Strabon
- Experimental evaluation
- Future Work and Conclusions

The “s” in stRDF and stSPARQL

[ISWC '12]

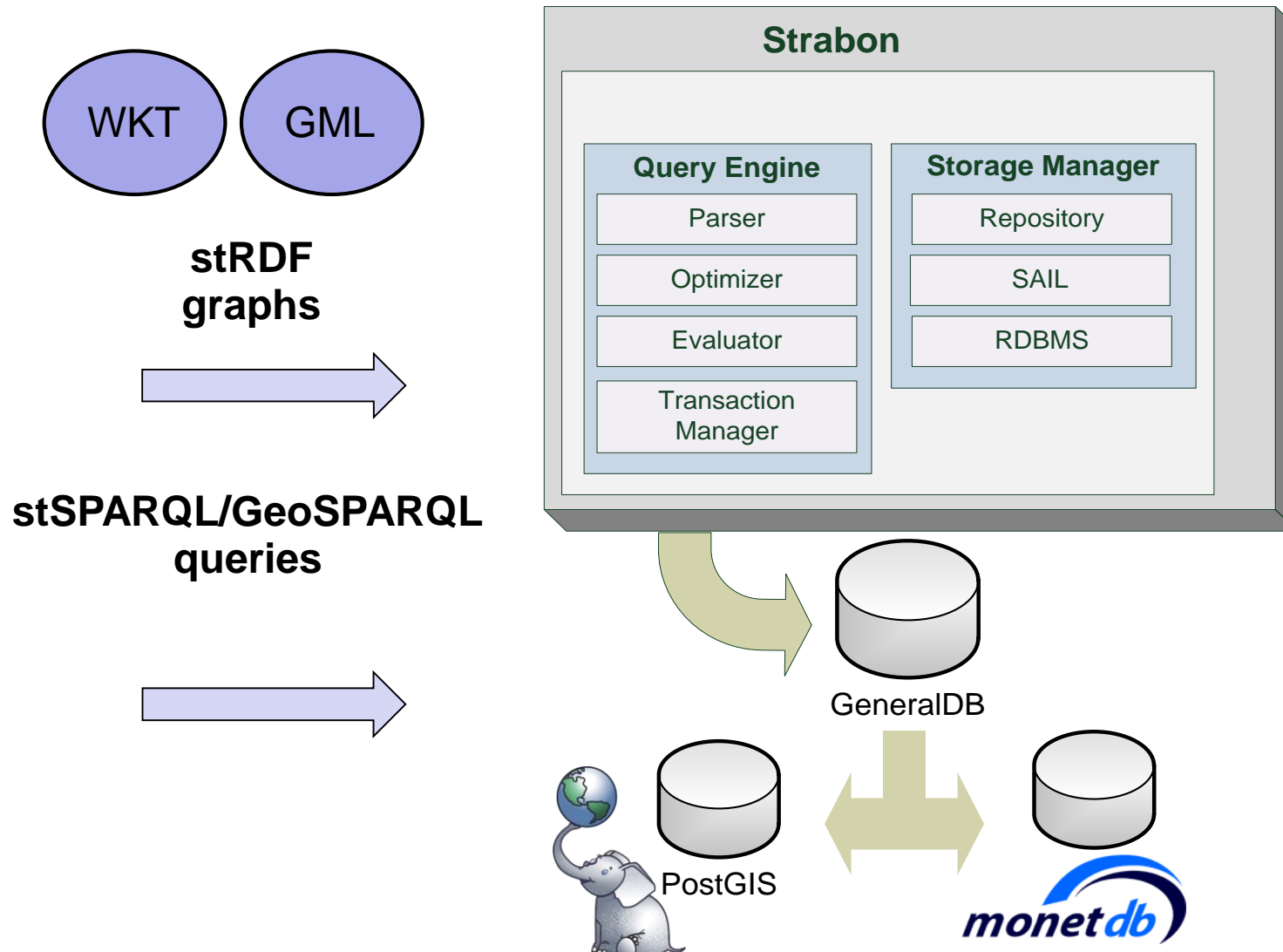
- **stRDF:** Extends RDF for representing geospatial data
 - **Objects in triples can be spatial literals** encoded in Well-Known Text/GML (OGC standards)



- **stSPARQL:** Extends SPARQL 1.1 with **spatial extension functions** defined in OGC standards for querying geospatial data.

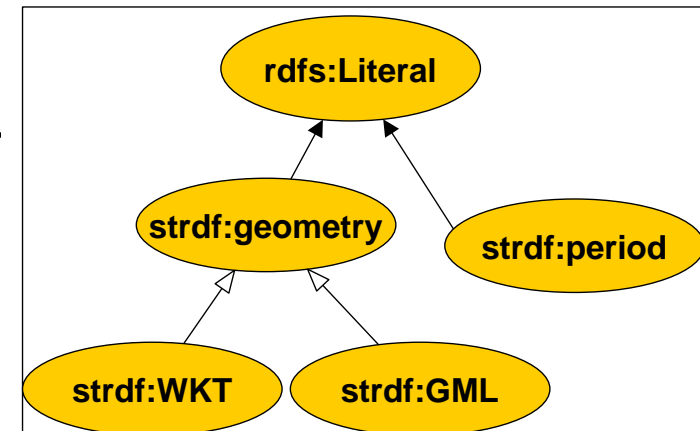
Strabon (<http://strabon.di.uoa.gr>)

[ISWC '12]



The "t" in stRDF

- **User-defined time:** already supported in RDF
- **Valid time:** Triples are annotated by their valid time
 - **Triples** are extended to **quads** (the last component is the valid time of the triple)
 - **Instants** represented using `xsd:dateTime` values
 - **Periods** represented using literals of the new datatype `strdf:period`
 - **Temporal Constants:** `NOW`, `UC` ("Until Changed")

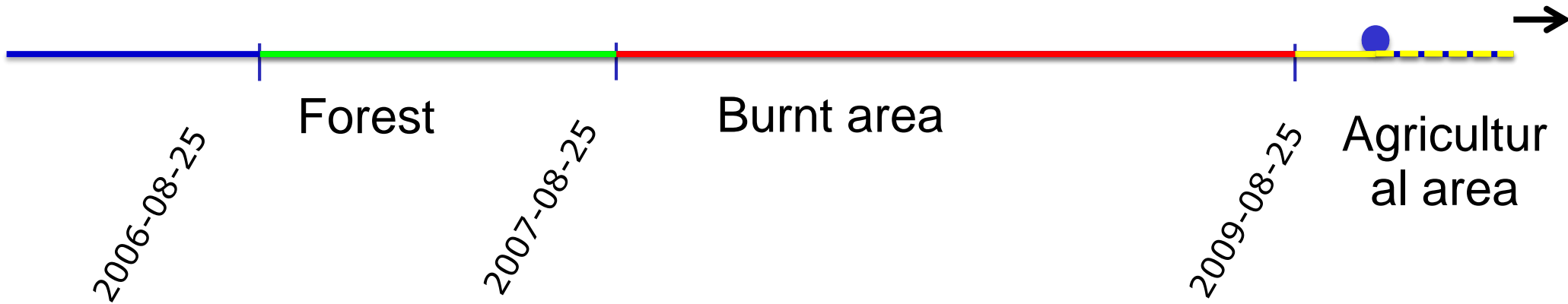


The “t” in stSPARQL

SPARQL 1.1 is extended as follows:

- **Triple patterns** extended to **quad patterns**
(the last component is a temporal term)
- **Temporal extension functions**
 - Allen's temporal relations (e.g., `before`, `meets`, `overlaps`)
 - Nine more functions that are syntactic sugar to the above
(e.g., `intersects`, `startsWithorAfter`, `finishesWithOrBefore`)
 - Interval-to-point relations (e.g., `during`, `before`, `after`)
 - Period constructors (e.g., `period_union`, `period_intersect`)
 - Temporal aggregates (e.g., `maximalPeriod`, `intersectAll`)
 - Temporal Updates
- Defined under strdf namespace: <http://strdf.di.uoa.gr/ontology>
- Complete stSPARQL reference:
<http://www.strabon.di.uoa.gr/stSPARQL>

Example



Example stRDF graph



```
clc:region1 rdf:type clc:Region .
```

```
clc:region1 strdf:hasGeometry  
"POLYGON((23.38 40.48, ... 23.38 40.48))"^^strdf:WKT .
```

```
clc:region1 clc:hasLandCover clc:Forest  
"[2006-08-25T11:00:00+02,2007-08-25T11:00:00+02)"^^strdf:period .
```

Quad

```
noa:ba1 rdf:type noa:BurntArea  
"[2007-08-25T11:00:00+02,2009-08-25T11:00:00+02)"^^strdf:period .
```

```
noa:ba1 strdf:hasGeometry  
"POLYGON((23.26 40.51, ... 23.26 0.51))"^^strdf:WKT.
```

```
clc:region1 clc:hasLandCover clc:AgriculturalArea  
"[2009-08-25T11:00:00+02, "UC)"^^strdf:period .
```

Temporal Constant

stSPARQL query (1/2)

- Find the **current** land cover of all areas in the dataset

```
SELECT ?clc
```

```
WHERE {
```

```
  ?R rdf:type clc:Region .
```

```
  ?R clc:hasLandCover ?clc ?t1 .
```

```
  FILTER(strdf:during ("NOW", ?t1))
```

```
}
```

Quad Pattern

Temporal constant

Temporal extension function

stSPARQL query (2/2)

Find all areas that were first covered by a forest, then got burnt, and then became agricultural areas.

```
SELECT ?R
```

```
WHERE {
```

```
  ?R rdf:type clc:Region .
```

```
  ?R strdf:hasGeometry ?RGEO .
```

```
  ?R clc:hasLandCover clc:Forest ?t1 .
```

```
  ?A rdf:type noa:BurntArea ?t2 .
```

```
  ?A strdf:hasGeometry ?BGEO .
```

```
  FILTER(strdf:after(?t2, ?t1)) .
```

```
  FILTER(strdf:contains( ?BGEO, ?RGEO)) .
```

```
  ?R clc:hasLandCover clc:AgriculturalArea ?t3 .
```

```
  FILTER(strdf:after(?t3, ?t2)) }
```

Temporal extension
function

Outline

- Introduction
- The “t” of stRDF and stSPARQL
- **Implementation in Strabon**
- Experimental evaluation
- Future Work and Conclusions

Translating quads to triples

- Example:

`clc:region1 clc:hasLandCover clc:Forest`

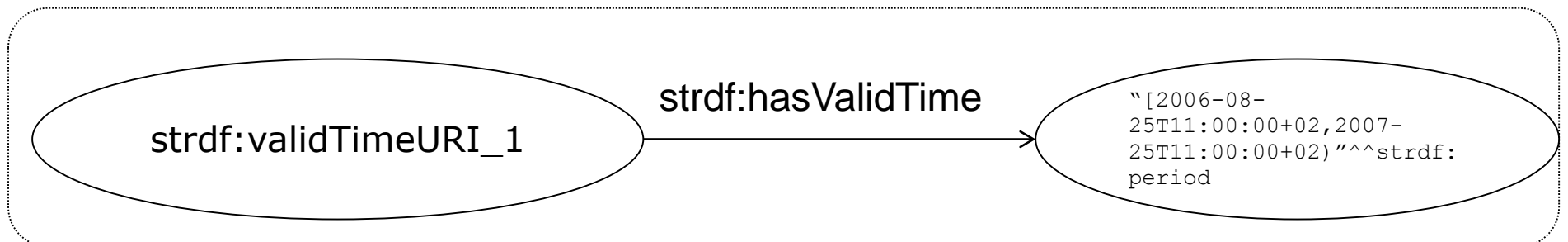
`"[2006-08-25T11:00:00+02,2007-25T11:00:00+02)""^^strdf:period.`



GRAPH strdf:validTimeURI_1



DEFAULT GRAPH



Translating quad patterns to triple patterns

Example

```
SELECT ?clc
```

```
WHERE {?R rdf:type clc:Region .
```

```
  ?R clc:hasLandCover ?clc ?t1 .
```

```
  FILTER(strdf:during ("NOW", ?t1))}
```



```
SELECT ?clc
```

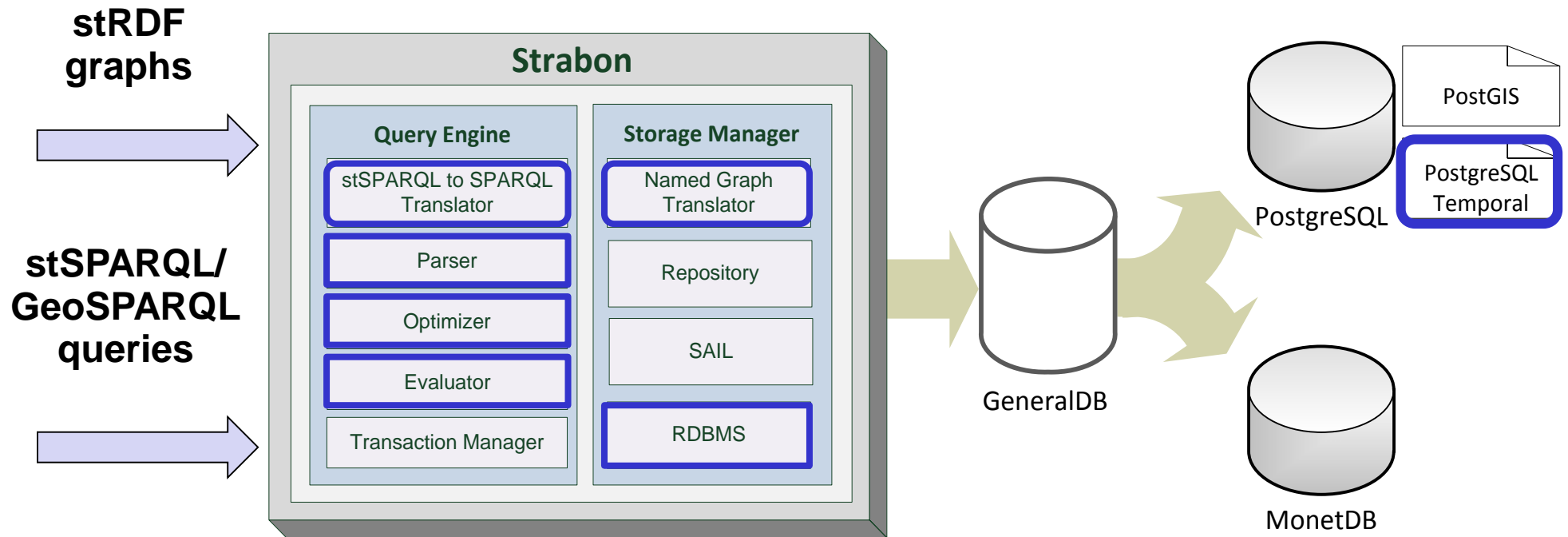
```
WHERE {?R rdf:type clc:Region .
```

```
  GRAPH ?g {?R clc:hasLandCover ?clc .}
```

```
  ?g strdf:hasValidTime ?t1 .
```

```
  FILTER(strdf:during ("NOW", ?t1))}
```

Strabon extended with valid time



Outline

- Introduction
- The “t” of stRDF and stSPARQL
- Implementation in Strabon
- **Experimental evaluation**
- Future Work and Conclusions

Competitor Systems

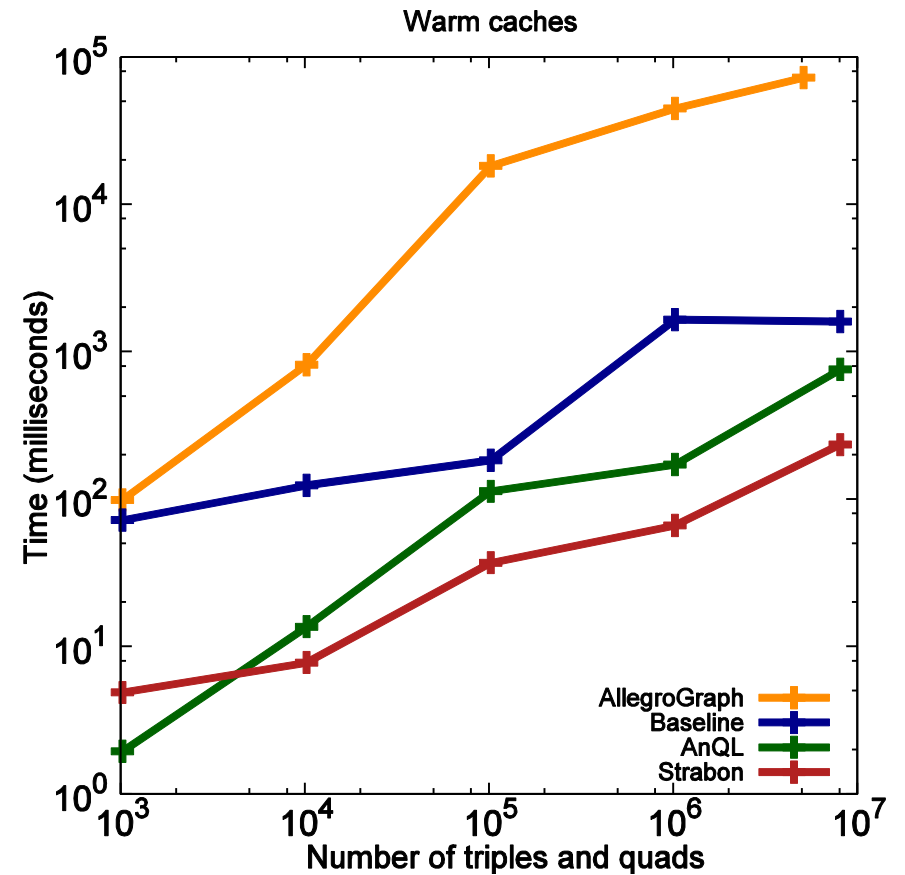
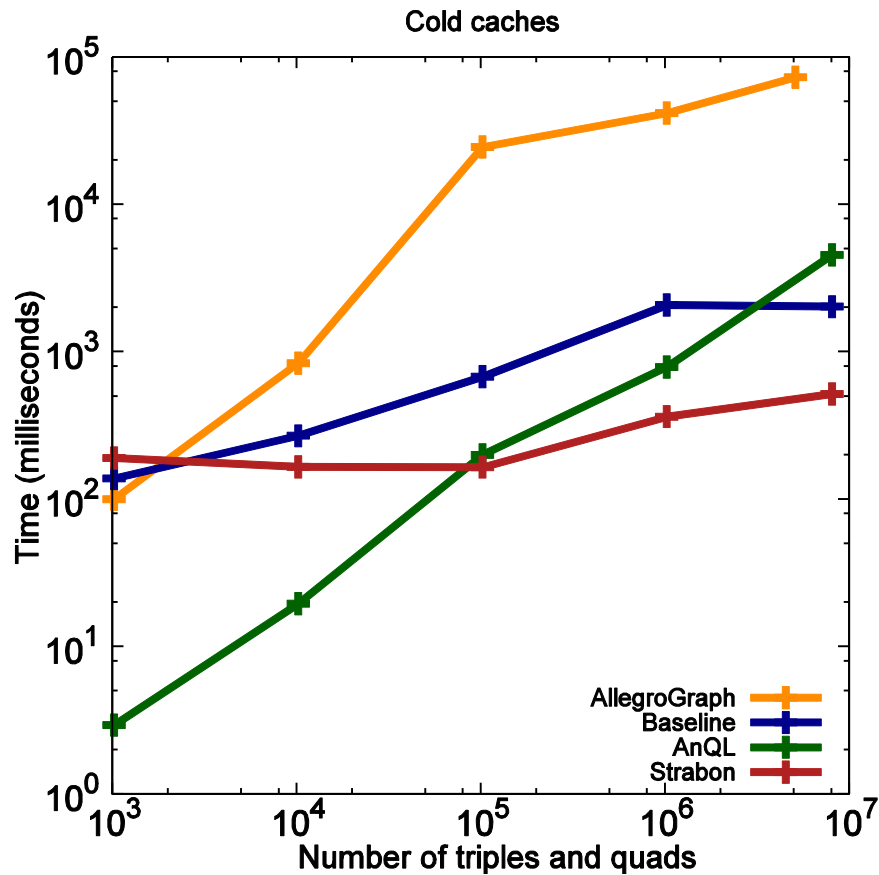
- Baseline implementation: [Sesame Native Store](#)
- Prolog-based implementation of [AnQL](#) (Lopes et al.)
- [AllegroGraph](#) (Free server edition v4.10)

Datasets

- GovTrack dataset
 - 8M triples
 - 42K periods
 - 294K instants

- Corine Land Cover changesets 2000-2006
 - 795K triples
 - 718K periods

Query response time as input dataset grows



Query response time with respect to query complexity – cold caches(sec)

System	Q1	Q2	Q3	Q4	Q5	Q6	Q7
Strabon	0.3	0.4	0.6	0.6	0.7	0.4	0.4
Baseline	0.9	2.1	2.2	2.3	2.3	0.4	0.4
AnQL	2.0	1.7	4.2	4.3	5.8	6.9	7.4
AllegroGraph	1.0	58.1	91.7	121.8	154.6	33.8	156.4
Results	12203	781	781	781	781	128	87

Query response time with respect to query complexity - warm caches (sec)

System	Q1	Q2	Q3	Q4	Q5	Q6	Q7
Strabon	0.01	0.07	0.9	0.1	0.1	0.04	0.04
Baseline	0.7	1.6	1.7	1.7	1.8	0.4	0.5
AnQL	0.2	0.2	0.2	0.3	0.9	0.4	0.5
AllegroGraph	1.0	56.3	90.8	123.5	154.6	157.8	158.0
Results	12203	781	781	781	781	128	87

Outline

- Introduction
- The “t” of stRDF and stSPARQL
- Implementation in Strabon
- Experimental evaluation
- **Future Work and Conclusions**

Conclusions and Future work

Conclusions

- We defined the valid time dimension of stSPARQL
- Strabon is a very rich spatio-temporal RDF store, offering efficient evaluation of temporal operators

Future work

- Experiments with even larger datasets, different storage schemes and query evaluation strategies
- Continue systematic evaluation of temporal RDF stores
- Representation and querying of moving objects

Thank you for your attention!

Questions?

- Strabon (<http://strabon.di.uoa.gr>)
 - Manolis Koubarakis, Kostis Kyzirakos, Manos Karpathiotakis, Charalampos Nikolaou, Giorgos Garbis, Konstantina Bereta, Kallirroï Dogani, Stella Giannakopoulou and Panayiotis Smeros.
 - Mercurial repository: <http://hg.strabon.di.uoa.gr>
 - Trac: <http://bug.strabon.di.uoa.gr>
 - Mailing list: <http://cgi.di.uoa.gr/~mailman/listinfo/strabon-users>
- Sextant: A web tool for browsing and mapping Linked Geospatial Data (<http://test.strabon.di.uoa.gr/sextant/>) (**visit our demo!**)
- TELEIOS EU Project (<http://www.earthobservatory.eu>) (**EU Project NW session**)

