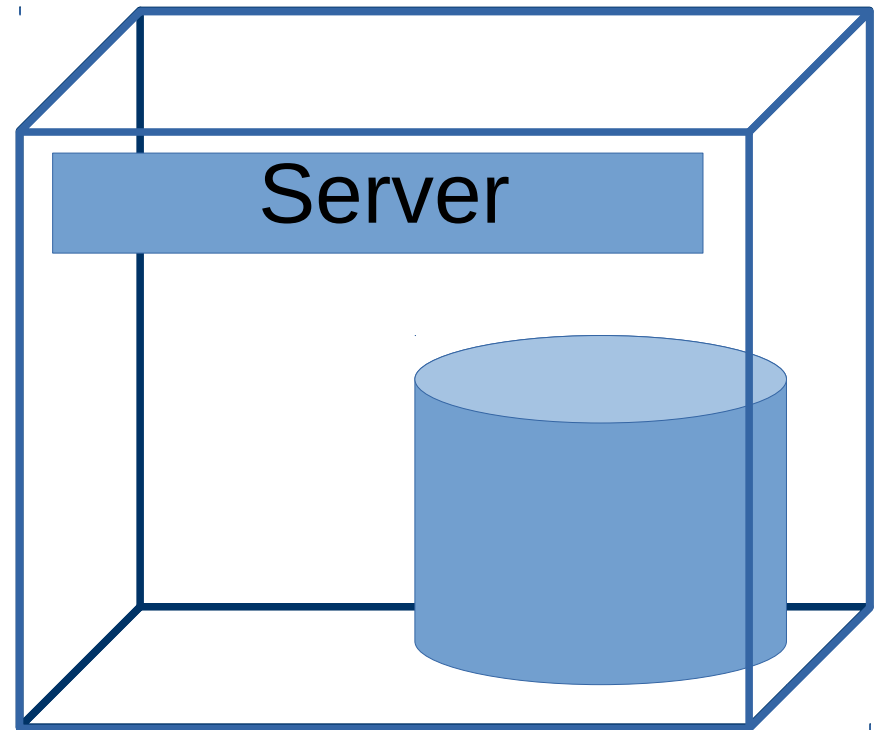


# Linked Data Fragments

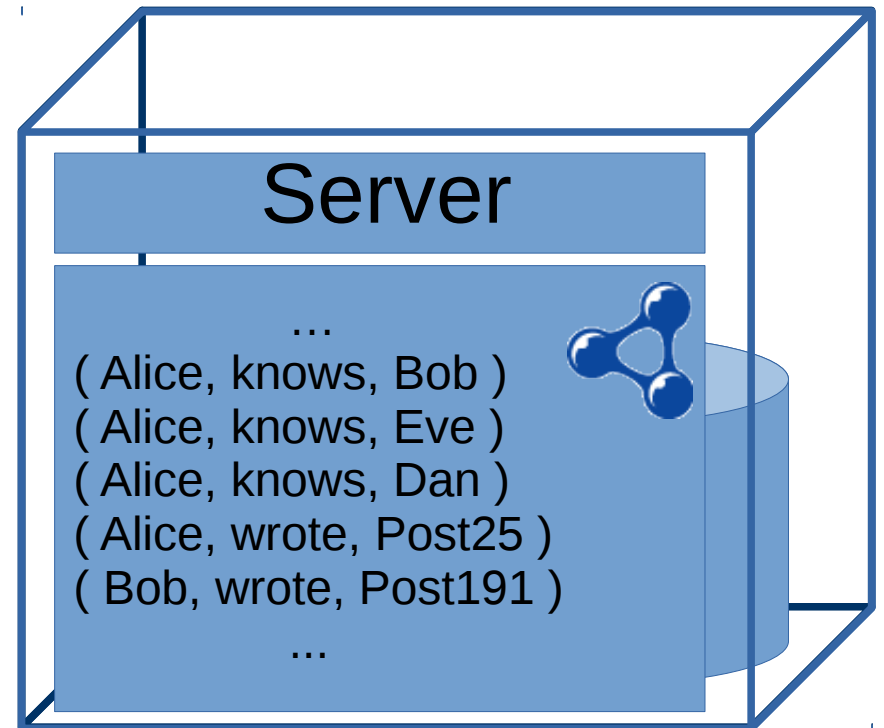
Exploring the Trade-Offs of Web Interfaces to  
Support Live Queries over (Semantic) Web Data

Olaf Hartig

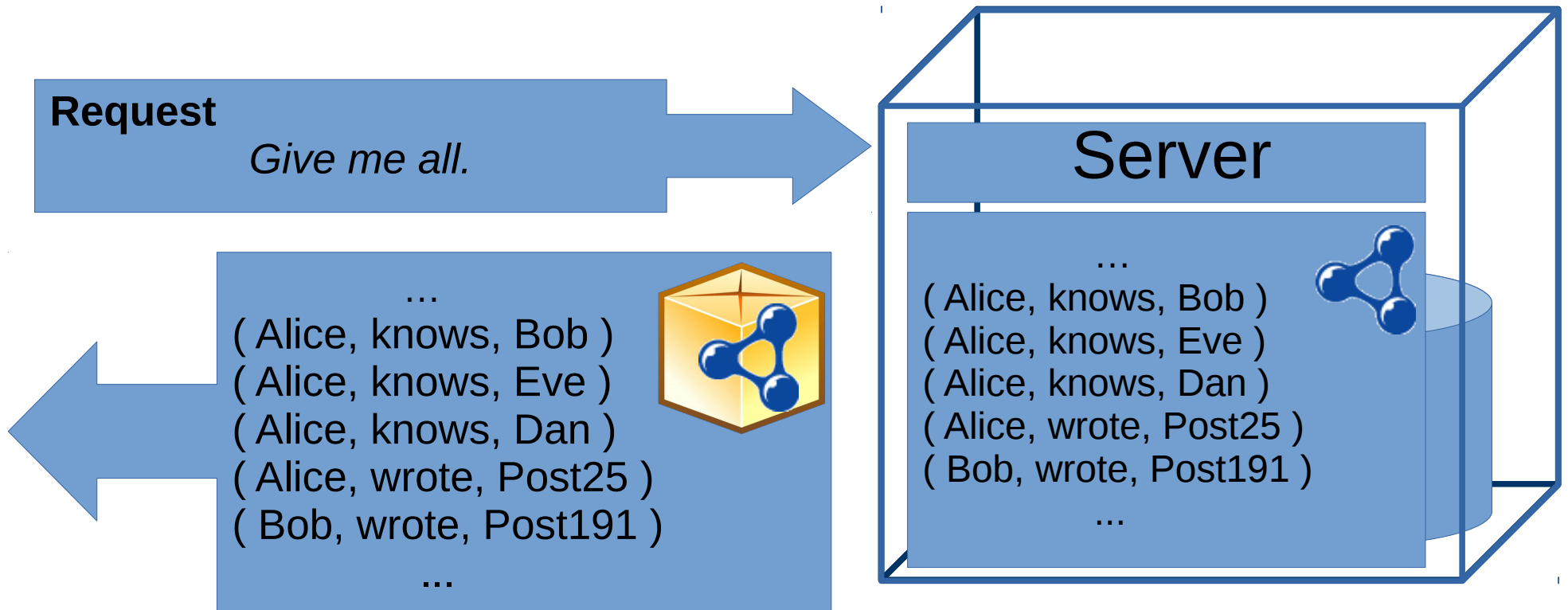
@olafhartig



# Semantic Web Solutions (So Far)

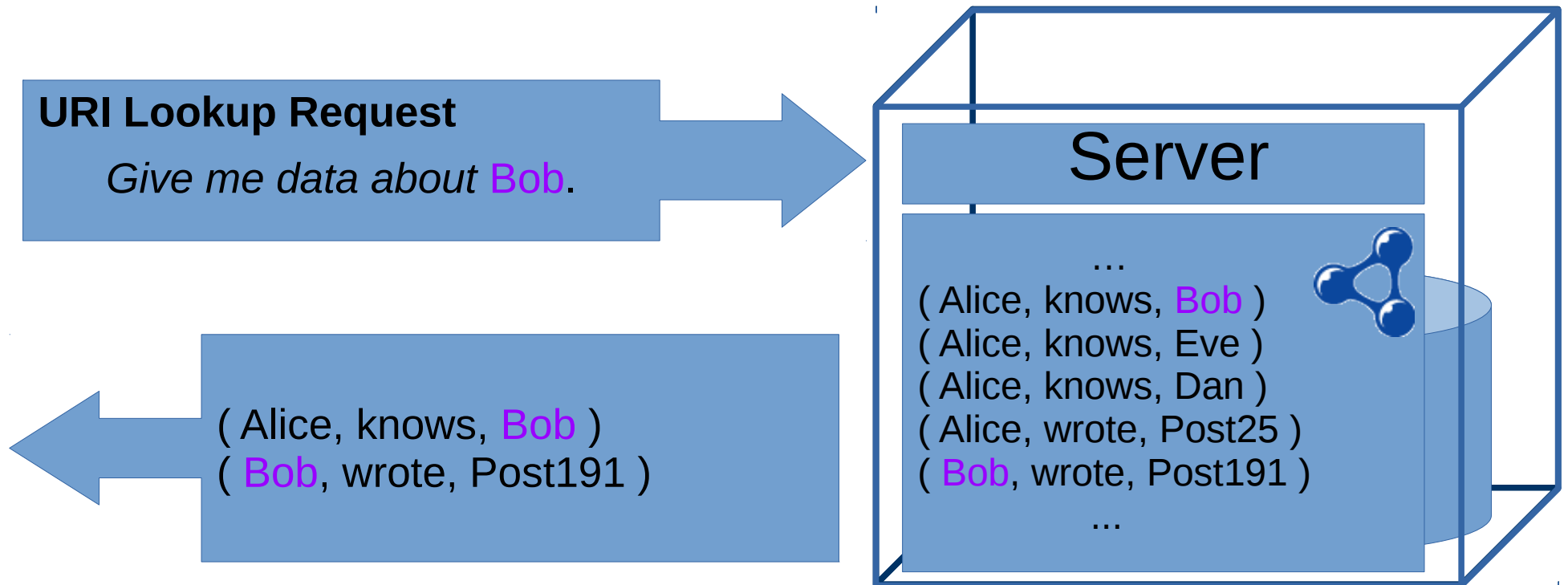


# Semantic Web Solutions (So Far)



RDF data  
dump

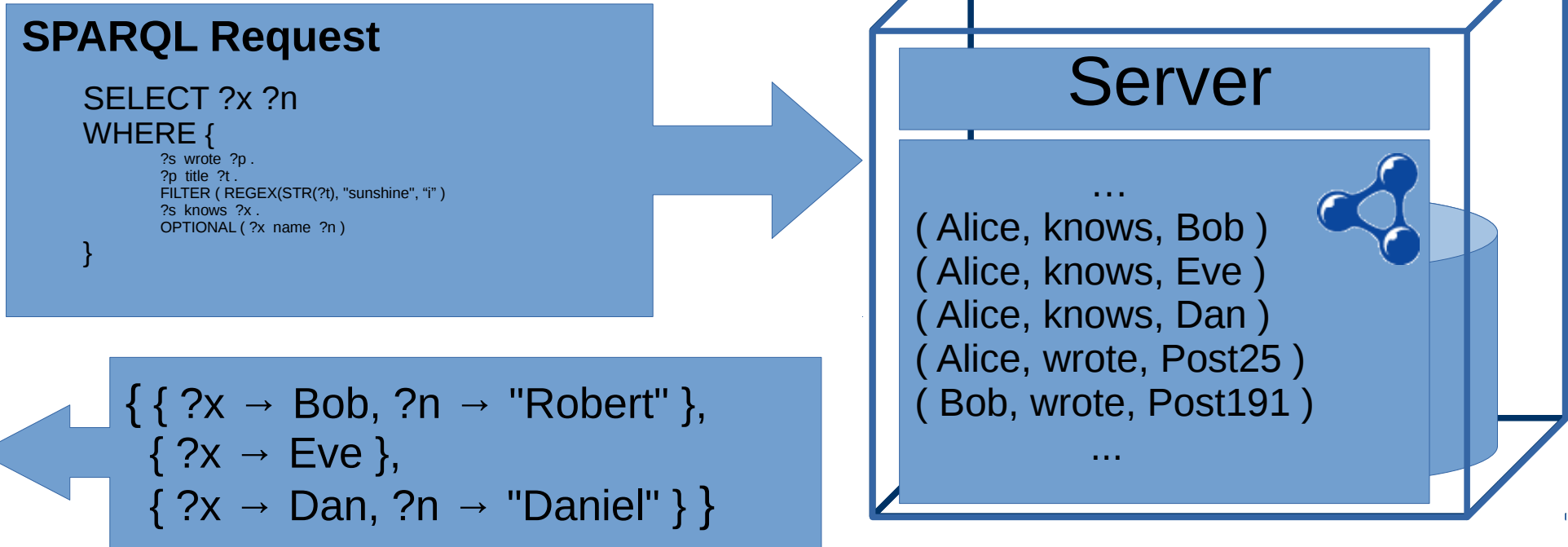
# Semantic Web Solutions (So Far)



RDF data  
dump

Linked Data  
documents

# Semantic Web Solutions (So Far)

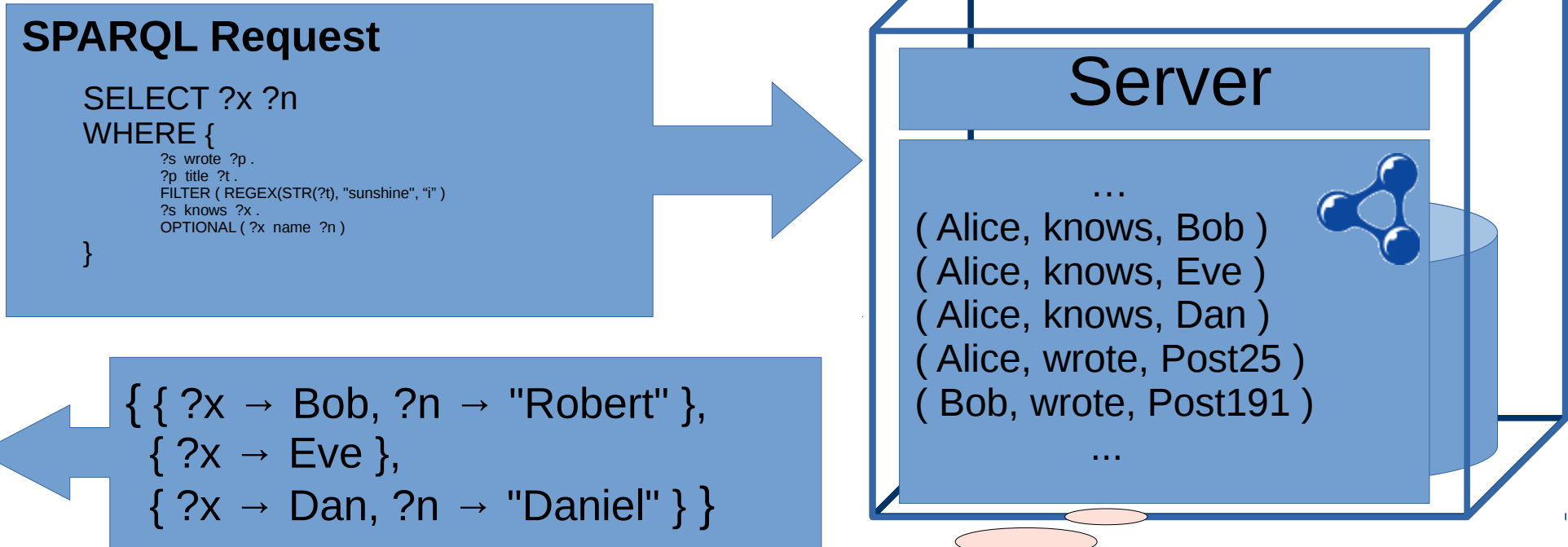


RDF data  
dump

Linked Data  
documents

SPARQL  
endpoint

# Semantic Web Solutions (So Far)



SPARQL endpoint

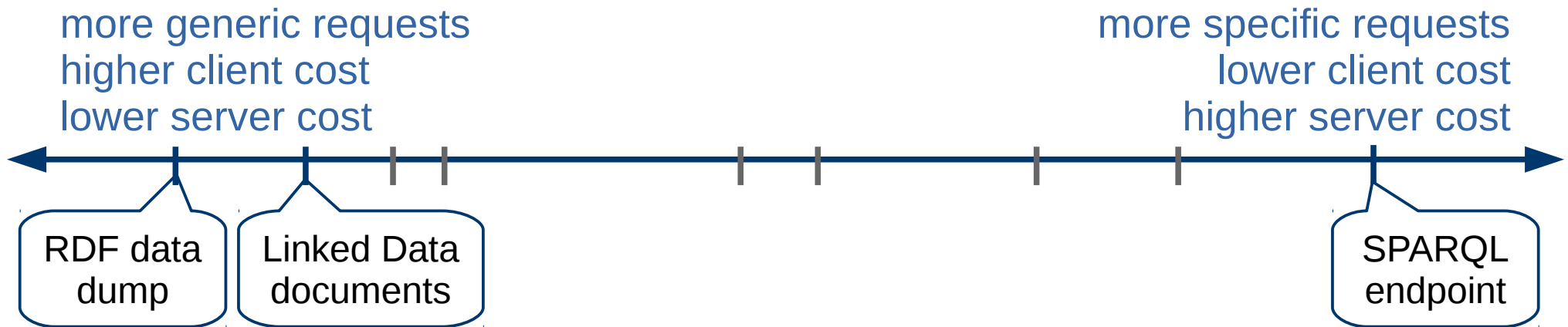
Out of 427 public SPARQL endpoints, **more than half** had **<95% availability**.<sup>1</sup>

→ not available for at least 1.5 days each month

<sup>1</sup> C. Buil Aranda, A. Hogan, J. Umbrich, et al.: *SPARQL Web-Querying Infrastructure: Ready for Action?* ISWC 2013.

# Linked Data Fragments<sup>1,2</sup>

- Whole spectrum of trade-offs exists between these extremes
- Explore this spectrum and find interesting sweet spots



<sup>1</sup> R. Verborgh, **O. Hartig**, B. De Meester, et al.: *Querying Datasets on the Web with High Availability*. ISWC 2014.

<sup>2</sup> R. Verborgh, M. Vander Sande, **O. Hartig**, et al.: *Triple Pattern Fragments: a Low-cost Knowledge Graph Interface for the Web*. In *Journal of Web Semantics* 37-38, 2016

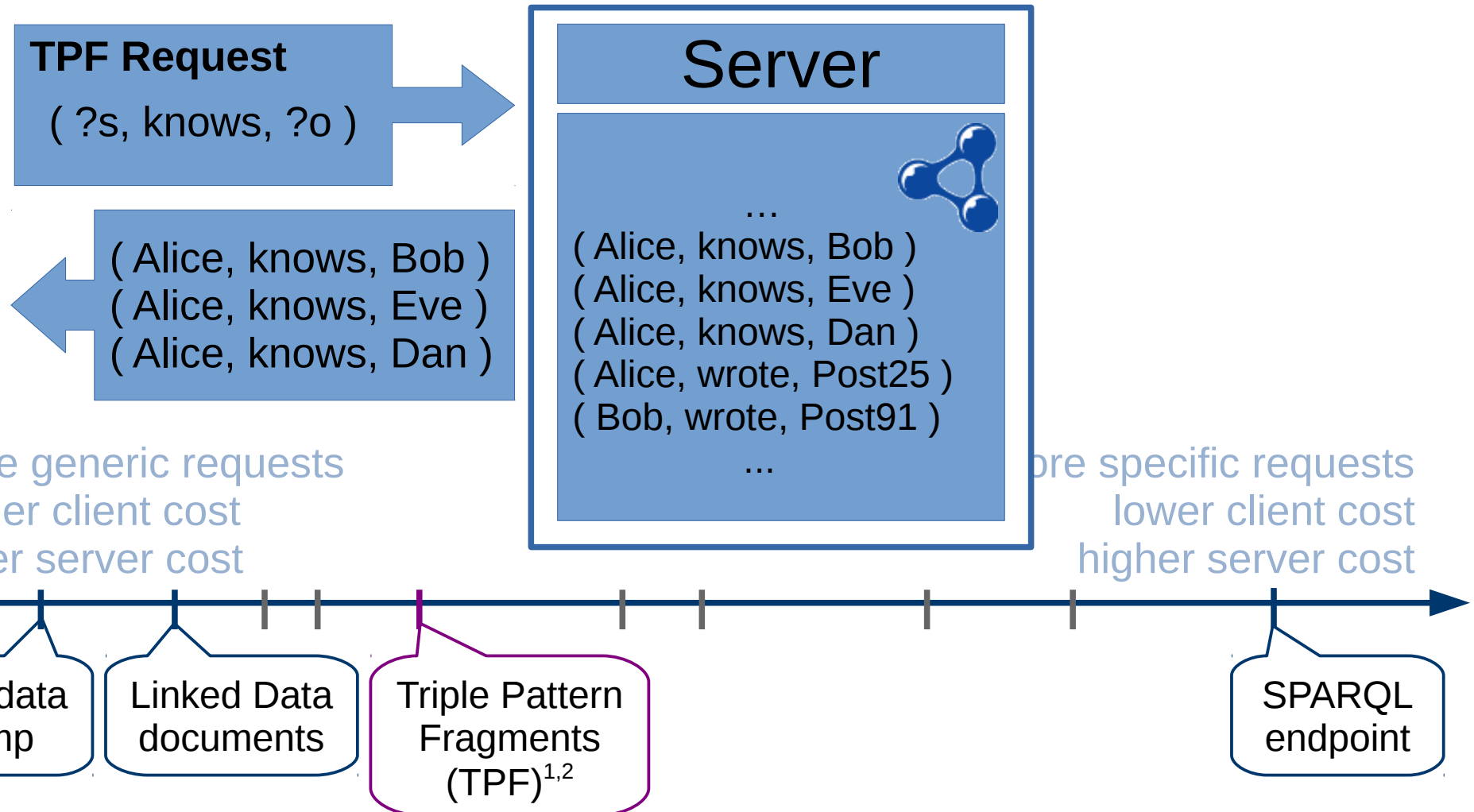


# Triple Pattern Fragments (TPF)<sup>1,2</sup>

<sup>1</sup> R. Verborgh, **O. Hartig**, B. De Meester, et al.: *Querying Datasets on the Web with High Availability*. ISWC 2014.

<sup>2</sup> R. Verborgh, M. Vander Sande, **O. Hartig**, et al.: *Triple Pattern Fragments: a Low-cost Knowledge Graph Interface for the Web*. In *Journal of Web Semantics* 37-38, 2016

# Triple Pattern Fragments (TPF)



<sup>1</sup> R. Verborgh, O. Hartig, B. De Meester, et al.: *Querying Datasets on the Web with High Availability*. ISWC 2014.

<sup>2</sup> R. Verborgh, M. Vander Sande, O. Hartig, et al.: *Triple Pattern Fragments: a Low-cost Knowledge Graph Interface for the Web*. In *Journal of Web Semantics* 37-38, 2016

# TPF-based Execution of SPARQL Queries

## SPARQL Query

```
SELECT ?y WHERE {  
  Alice knows ?x .  
  ?x wrote ?y }
```

( Alice, knows, Bob )  
( Alice, knows, Eve )  
( Alice, knows, Dan )

( Alice, wrote, Post25 )  
( Bob, wrote, Post91 )  
...

## TPF Request

( Alice, knows, ?x )

( Alice, knows, Bob )  
( Alice, knows, Eve )  
( Alice, knows, Dan )

## TPF Request

( ?x, wrote, ?y )

( Alice, wrote, Post25 )  
( Bob, wrote, Post91 )  
...

## Server

...

( Alice, knows, Bob )  
( Alice, knows, Eve )  
( Alice, knows, Dan )  
( Alice, wrote, Post25 )  
( Bob, wrote, Post91 )  
...

<sup>1</sup> R. Verborgh, O. Hartig, B. De Meester, et al.: *Querying Datasets on the Web with High Availability*. ISWC 2014.

<sup>2</sup> R. Verborgh, M. Vander Sande, O. Hartig, et al.: *Triple Pattern Fragments: a Low-cost Knowledge Graph Interface for the Web*. In *Journal of Web Semantics* 37-38, 2016

# TPF-based Execution of SPARQL Queries

## SPARQL Query

```
SELECT ?y WHERE {
  Alice knows ?x .
  ?x wrote ?y }
```

```
( Alice, knows, Bob )
( Alice, knows, Eve )
( Alice, knows, Dan )
```

## TPF Request

```
( Alice, knows, ?x )
```

```
( Alice, knows, Bob )
( Alice, knows, Eve )
( Alice, knows, Dan )
```

## TPF Request

```
( Bob, wrote, ?y )
```

```
( Bob, wrote, Post91 )
```

## TPF Request

```
( Eve, wrote, ?y )
```

*empty*

## Server

```
...
( Alice, knows, Bob )
( Alice, knows, Eve )
( Alice, knows, Dan )
( Alice, wrote, Post25 )
( Bob, wrote, Post91 )
...
```

<sup>1</sup> R. Verborgh, O. Hartig, B. De

Web with High Availability. ISWC 2014.

<sup>2</sup> R. Verborgh, M. Vander Sande, O. Hartig, et al.: *Triple Pattern Fragments: a Low-cost Knowledge Graph Interface for the Web*. In Journal of Web Semantics 37-38, 2016

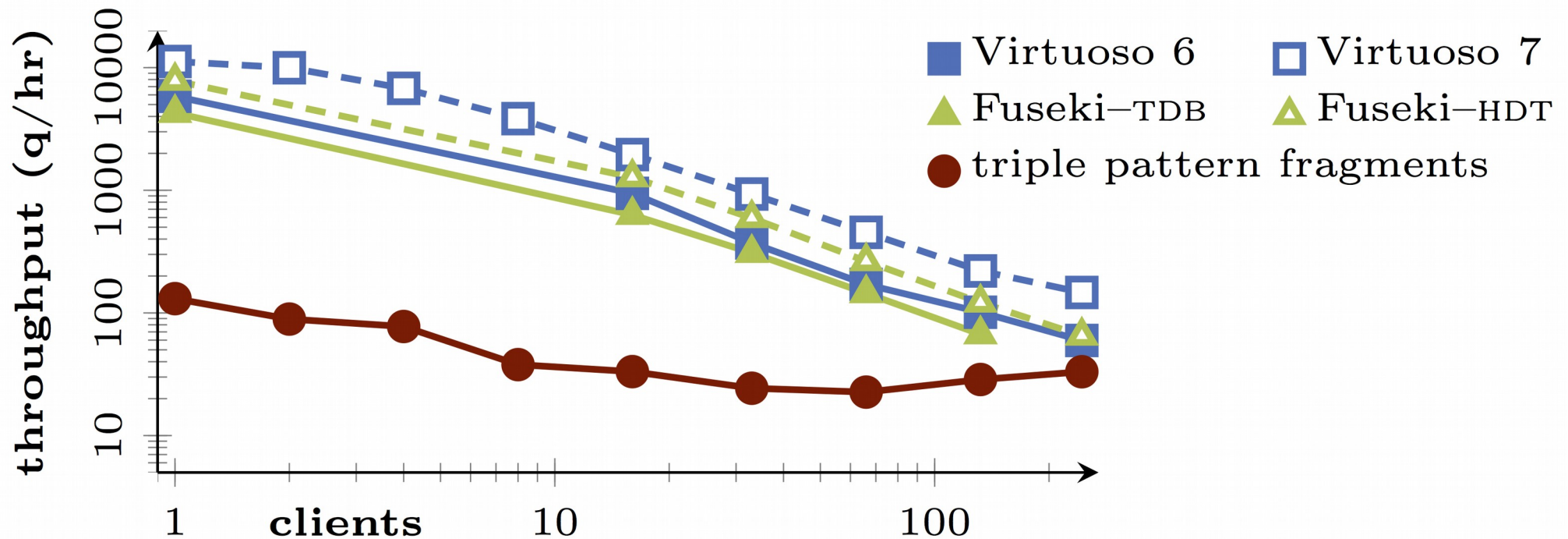
# Experimental Setup

- Berlin SPARQL Benchmark
  - Synthetic benchmark
  - 100M triples
- Amazon EC2 machines
  - 1 server (used either as SPARQL endpoint or as TPF server)
  - 1 cache
  - 1–240 clients

<sup>1</sup> R. Verborgh, **O. Hartig**, B. De Meester, et al.: *Querying Datasets on the Web with High Availability*. ISWC 2014.

<sup>2</sup> R. Verborgh, M. Vander Sande, **O. Hartig**, et al.: *Triple Pattern Fragments: a Low-cost Knowledge Graph Interface for the Web*. In *Journal of Web Semantics* 37-38, 2016

# Throughput (normalized by # of clients)

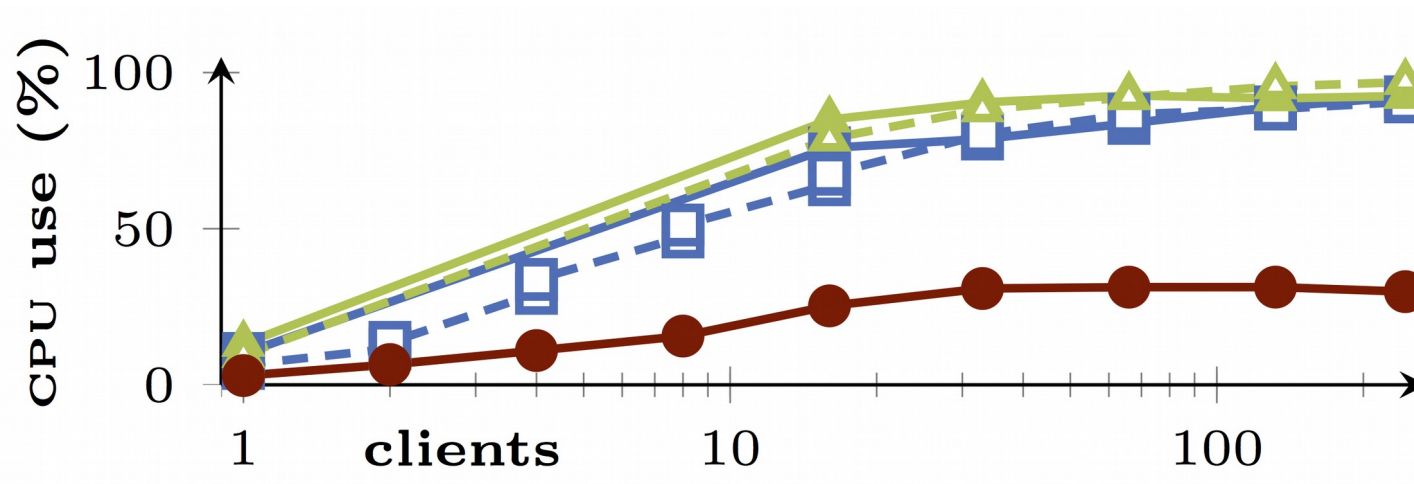


Observation: query throughput of TPF is lower but resilient to high client numbers

<sup>1</sup> R. Verborgh, O. Hartig, B. De Meester, et al.: *Querying Datasets on the Web with High Availability*. ISWC 2014.

<sup>2</sup> R. Verborgh, M. Vander Sande, O. Hartig, et al.: *Triple Pattern Fragments: a Low-cost Knowledge Graph Interface for the Web*. In *Journal of Web Semantics* 37-38, 2016

# Server-Side CPU Load

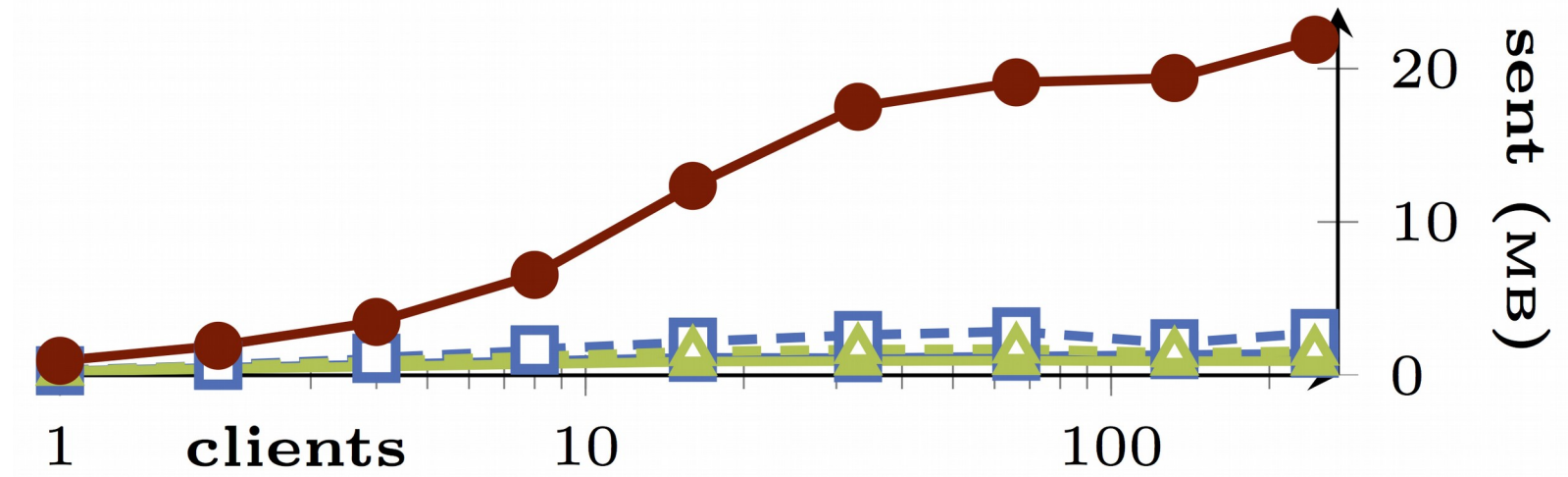


Observation: TPF server uses much less CPU

<sup>1</sup> R. Verborgh, **O. Hartig**, B. De Meester, et al.: *Querying Datasets on the Web with High Availability*. ISWC 2014.

<sup>2</sup> R. Verborgh, M. Vander Sande, **O. Hartig**, et al.: *Triple Pattern Fragments: a Low-cost Knowledge Graph Interface for the Web*. In *Journal of Web Semantics* 37-38, 2016

# Cache Traffic (TPF vs. SPARQL Endpoints)



Observation: caching is significantly more effective  
(because clients reuse fragments for queries)

<sup>1</sup> R. Verborgh, **O. Hartig**, B. De Meester, et al.: *Querying Datasets on the Web with High Availability*. ISWC 2014.

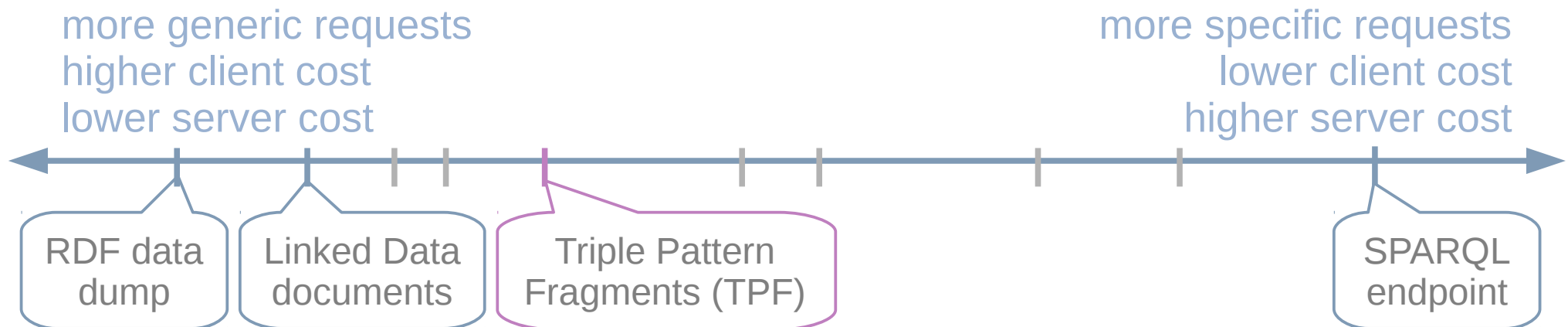
<sup>2</sup> R. Verborgh, M. Vander Sande, **O. Hartig**, et al.: *Triple Pattern Fragments: a Low-cost Knowledge Graph Interface for the Web*. In *Journal of Web Semantics* 37-38, 2016



# Summary of Experimental Results

Compared to SPARQL endpoints, query throughput is lower but ...

- ...resilient to high client numbers
- ...server-side load is much smaller and more regular (which allows for a higher availability, in particular on small, less expensive servers!)
- ...HTTP caching is significantly more effective



<sup>1</sup> R. Verborgh, **O. Hartig**, B. De Meester, et al.: *Querying Datasets on the Web with High Availability*. ISWC 2014.

<sup>2</sup> R. Verborgh, M. Vander Sande, **O. Hartig**, et al.: *Triple Pattern Fragments: a Low-cost Knowledge Graph Interface for the Web*. In *Journal of Web Semantics* 37-38, 2016

[www.liu.se](http://www.liu.se)