



Processing and Sharing Data Streams in Grid-Based P2P Infrastructures

Richard Kuntschke, Tobias Scholl, Bernhard Stegmaier, Alfons Kemper, and Angelika Reiser
Lehrstuhl Informatik III: Datenbanksysteme
Technische Universität München – Fakultät für Informatik

Data Basics

Data Model:

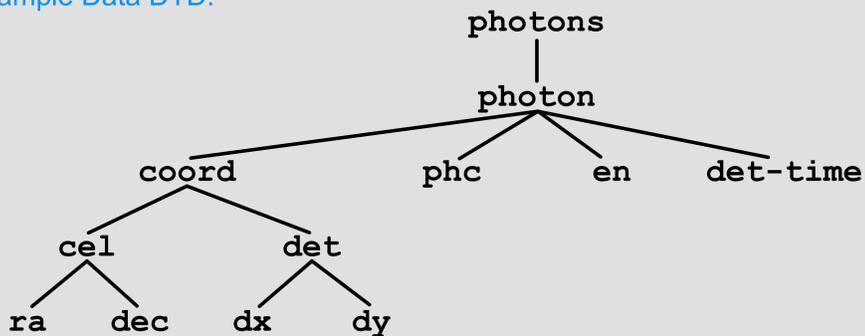
- XML data
- DTDs for data description

Example Data:

- Real astrophysical data
- ROSAT All Sky Survey (RASS)
- Collection of photon data



Example Data DTD:



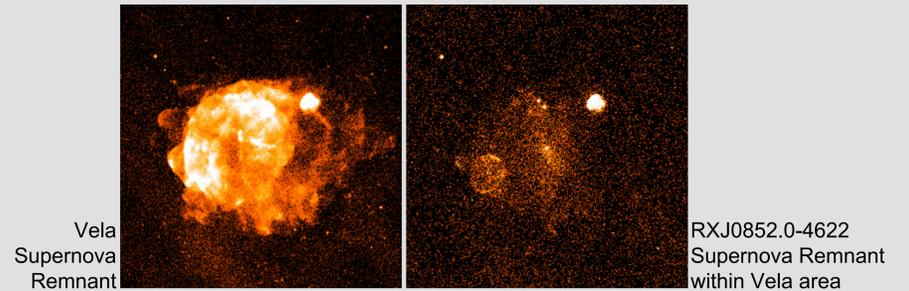
Query Basics

Query Language:

- Windowed XQuery (WXQuery)
- Fragment of XQuery
- Augmented with support for data windows (time-based and element-based), e.g., for window-based aggregation

Query Engine:

- *FluX* query engine (presented at VLDB 2004)
- Scalable evaluation of (W)XQueries on streaming data
- Uses schema information to reduce buffer consumption



Query 1

```

<photons>
  {for $p in stream("stream-0")/photons/photon
   where $p/coord/cel/ra >= 120.0
     and $p/coord/cel/ra <= 138.0
     and $p/coord/cel/dec >= -49.0
     and $p/coord/cel/dec <= -40.0
   return
    <vela_photon>
      {$p/coord/cel/ra} {$p/coord/cel/dec}
      {$p/phc} {$p/en} {$p/det-time}
    </vela_photon>}
</photons>
  
```

- Selects the area of the Vela supernova remnant
- Returns information on photon coordinates, detector pulse, photon energy, and detection time

Query 2

```

<photons>
  {for $p in stream("stream-0")/photons/photon
   where $p/en >= 1.3
     and $p/coord/cel/ra >= 130.5
     and $p/coord/cel/ra <= 135.5
     and $p/coord/cel/dec >= -48.0
     and $p/coord/cel/dec <= -45.0
   return
    <rxj_photon>
      {$p/coord/cel/ra} {$p/coord/cel/dec}
      {$p/en} {$p/det-time}
    </rxj_photon>}
</photons>
  
```

- Selects highly energetic photons in the area of the RXJ0852.0-4622 supernova remnant
- Can reuse result data stream of Query 1 as input

Query 3

```

<photons>
  {for $w in stream("stream-0")/photons/photon
   [en >= 1.3
    and coord/cel/ra >= 130.5
    and coord/cel/ra <= 135.5
    and coord/cel/dec >= -48.0
    and coord/cel/dec <= -45.0]
   |count 20 step 10|
   let $a := avg($w/photon/en)
   return
    <avg_en>
      {$a}
    </avg_en>}
</photons>
  
```

- Computes the average energy of 20 successive highly energetic photons in steps of 10 photons within the RXJ0852.0-4622 area
- Can reuse result data stream of Query 1 or 2 as input

Query 4

```

<photons>
  {for $w in stream("stream-0")/photons/photon
   [en >= 1.3
    and coord/cel/ra >= 130.5
    and coord/cel/ra <= 135.5
    and coord/cel/dec >= -48.0
    and coord/cel/dec <= -45.0]
   |count 60 step 40|
   let $a := avg($w/photon/en)
   return
    <avg_en>
      {$a}
    </avg_en>}
</photons>
  
```

- Computes the average energy of 60 successive highly energetic photons in steps of 40 photons within the RXJ0852.0-4622 area
- Can reuse result data stream of Query 1, 2, or 3 as input