

FluXQuery

An Optimizing XQuery Processor for Streaming XML Data

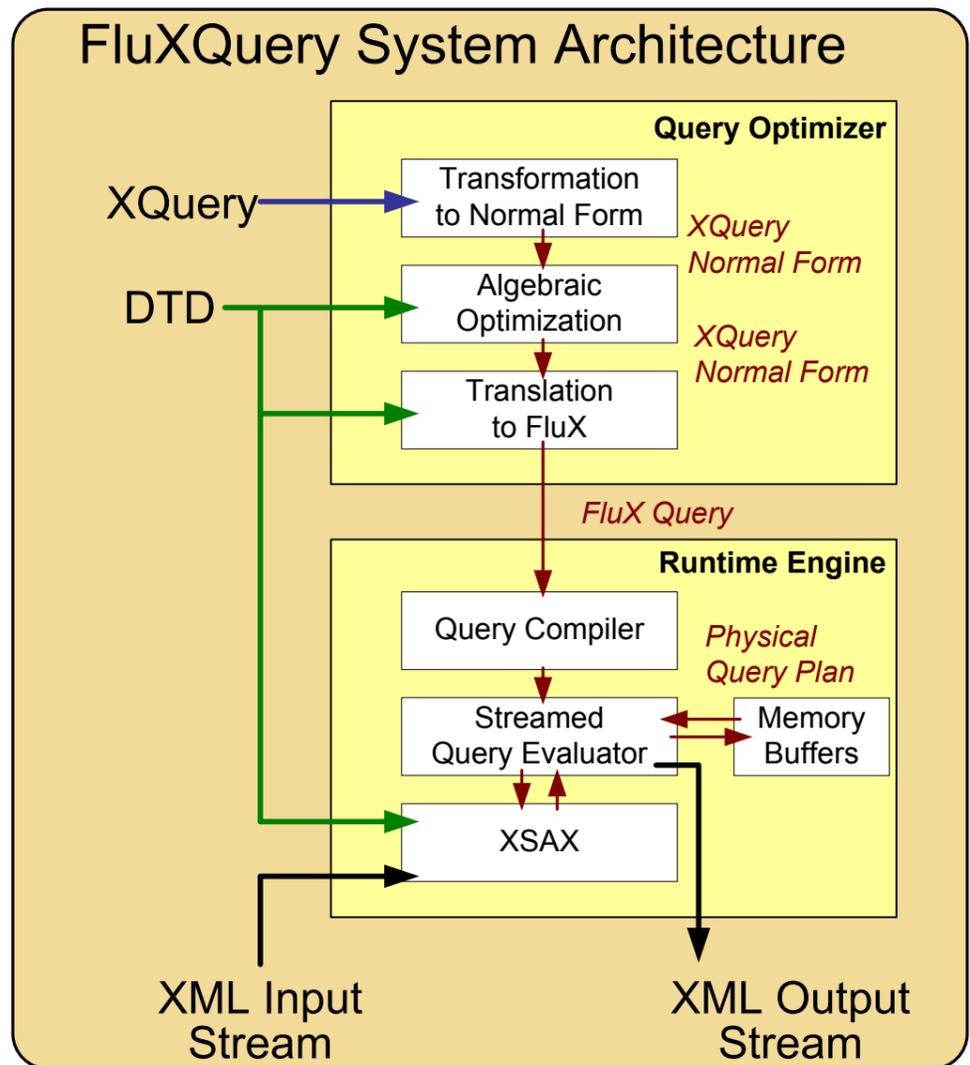
The FluXQuery engine supports

- a powerful fragment of XQuery:
 - nested for-loops and
 - joins
- event-based query processing
- conscious handling of main memory buffers
- algebraic optimization based on schema information

Christoph Koch
Stefanie Scherzinger
Nicole Schweikardt
Bernhard Stegmaier

Technische Universität Wien
Technische Universität Wien
Humboldt Universität zu Berlin
Technische Universität München

FluXQuery System Architecture



```
<results>
{ for $b in $ROOT/bib/book
return
  <result>
    { $b/title } { $b/author }
  </result>
}
</results>
```

XQuery

```
<results>
{ for $bib in $ROOT/bib return
  { for $b in $bib/book return
    <result>
      { for $t in $b/title return {$t} }
      { for $a in $b/author return {$a} }
    </result>
  }
}
</results>
```

XQuery Normal Form

```
<!ELEMENT book ((title|author)*,year)>
→ no order among titles and authors
```

DTD

```
<!ELEMENT book (title,author*,year)>
→ order constraint: title <book author
```

DTD

```
<results>
{ ps $ROOT: on bib as $bib return
  { ps $bib: on book as $b return
    <result>
      { ps $b:
        on title as $t return {$t};
        on-first past(title, author) return
          { for $a in $b/author return {$a} }
      }
    </result>
  }
}
</results>
```

buffer authors

FluX
Queries

```
<results>
{ ps $ROOT: on bib as $bib return
  { ps $bib: on book as $b return
    <result>
      { ps $b:
        on title as $t return {$t};
        on author as $a return {$a};
      }
    </result>
  }
}
</results>
```

no buffering

Reference: C. Koch, S. Scherzinger, N. Schweikardt, and B. Stegmaier. "Schema-based Scheduling of Event Processors and Buffer Minimization for Queries on Structured Data Streams". In *Proc. VLDB 2004*.