



Publishing and Querying the Histories of Archived Relational Databases in XML



Fusheng Wang and Carlo Zaniolo Computer Science Department, UCLA

Publish the History of Relational Databases as XML Documents

- Publish the **history** of relational databases as XML documents:
 - ◆ Natural ways to represent such history in XML
 - ◆ Historical queries can be expressed in XQuery as is—no extensions to the data model or query language required for temporal queries
 - ◆ Approach amenable to efficient implementation: query and storage efficiency of alternative approaches
- Gain: Temporal applications are very important and are not supported well by current databases

History of Tables Transaction-Time Relational Tables

- Timestamped tuple snapshots

name	empno	salary	title	deptno	DOB	start	end
Bob	10003	60000	Engineer	d01	1945-04-09	1995-01-01	1995-05-31
Bob	10003	70000	Engineer	d01	1945-04-09	1995-06-01	1995-09-30
Bob	10003	70000	Sr Engineer	d02	1945-04-09	1995-10-01	1996-01-31
Bob	10003	70000	Tech Leader	d02	1945-04-09	1996-02-01	1996-12-31

- Temporally grouped history of employees

name	empno	salary	title	deptno	DOB
Bob	10003	60000	Engineer	d01	1945-04-09
		1995-01-01:1996-05-31	1995-01-01:1995-09-30	1995-01-01:1995-09-30	
		70000	Sr Engineer	d02	
		1995-06-01:1996-12-31	1995-10-01:1996-12-31	1995-10-01:1996-12-31	

XML Representation of Database History a Temporally Grouped Data Model

```

<employees tstart="1995-01-01" tend="1996-12-31">
  <employee tstart="1995-01-01" tend="1996-12-31">
    <empno tstart="1995-01-01" tend="1996-12-31">10003</empno>
    <name tstart="1995-01-01" tend="1996-12-31">Bob</name>
    <salary tstart="1995-01-01" tend="1995-05-31">60000</salary>
    <salary tstart="1995-06-01" tend="1996-12-31">70000</salary>
    <title tstart="1995-01-01" tend="1995-09-30">Engineer</title>
    <title tstart="1995-10-01" tend="1996-01-31">Sr Engineer</title>
    <title tstart="1996-02-01" tend="1996-12-31">Tech Leader</title>
    <dept tstart="1995-01-01" tend="1995-09-30">QA</dept>
    <dept tstart="1995-10-01" tend="1996-12-31">RD</dept>
    <DOB tstart="1995-01-01" tend="1996-12-31">1945-04-09</DOB>
  </employee>
<!-- More... -->
</employees>

```

Temporal Queries with XQuery

- Temporal projection: retrieve the salary history of "Bob":

```

element salary_history {
  for $s in document("employees.xml")/
    employees/employee[fname="Bob"]/salary
  return $s }

```

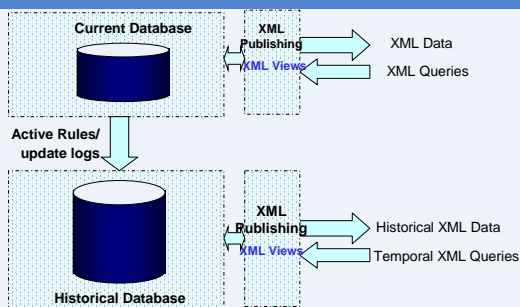
- Snapshot queries: retrieve the departments on 1996-01-31:

```

for $d in document("depts.xml")/depts/dept
[ tstart(.) <= "1996-01-31" and tend(.) >= "1996-01-31" ]
let $n := $d/name[tstart(.) <= "1996-01-31" and tend(.) >= "1996-01-31"]
let $m := $d/manager[tstart(.) <= "1996-01-31" and tend(.) >= "1996-01-31"]
return ( element dept{ $n, $m } )

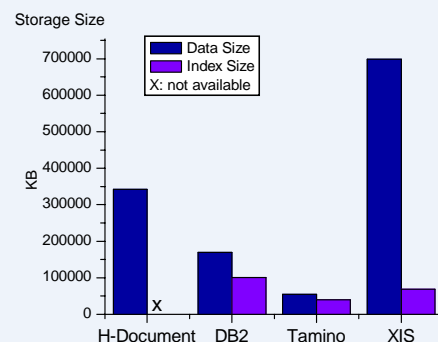
```

Historical Database Architecture



- Store history into RDBMS: history shredded into tables
- Store history into native XML databases

Performance Comparisons: Storage Size



Query Performance Comparisons

Query Performance of DB2 and Tamino:

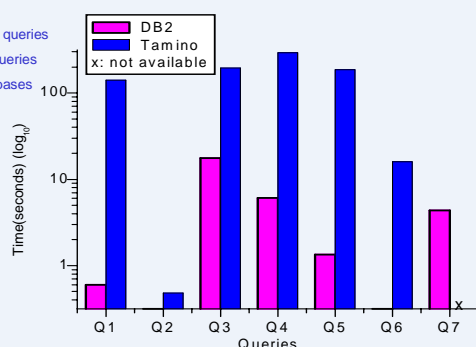
Q2: history query

Q4,Q6: snapshot queries

Q3,Q5: interval queries

Q1: scan of databases

Q7: join



Future Work

- Segment-based archiving scheme: clustering by segno into a single table on ATLaS
- Segment-based scheme on native XML database
- Mapping XQuery to SQL based on our schema
- Mapping user-defined functions in XQuery to SQL
- Compression support
- SQL/XML publishing functions support on ATLaS
- XPath support on ATLaS