

# Pathfinder: XQuery—The Relational Way

Jens Teubner<sup>1</sup> Peter Boncz<sup>2</sup> Torsten Grust<sup>1</sup>  
Maurice van Keulen<sup>3</sup> Stefan Manegold<sup>2</sup> Jan Rittinger<sup>2,4</sup>

<sup>1</sup>Technische Universität München, Germany

<sup>2</sup>CWI Amsterdam, The Netherlands

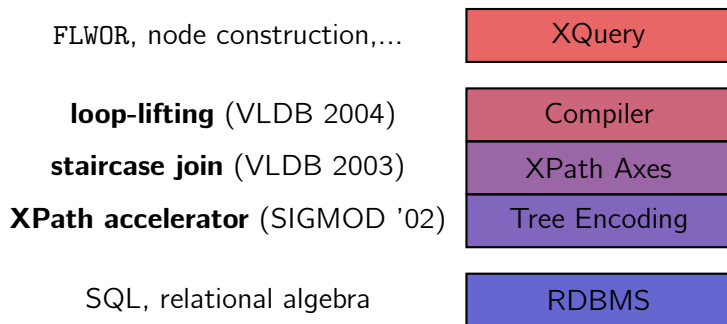
<sup>3</sup>University of Twente, The Netherlands

<sup>4</sup>University of Konstanz, Germany



Universität Konstanz

Pathfinder assembles relational XQuery processing techniques into a purely relational XQuery processing stack.



## We provide full XQuery support.

literals	42, "foo", (), ...
arithmetics	$e_1 + e_2$ , $e_1 - e_2$ , ...
built-in functions	fn:sum( $e$ ), fn:count( $e$ ), fn:doc( $uri$ ), ...
variable bindings	let $\$v := e_1$ return $e_2$
iteration	for $\$v$ at $\$p$ in $e_1$ return $e_2$
conditionals	if $p$ then $e_1$ else $e_2$
sequence construction	$e_1$ , $e_2$
function calls	$f(e_1, e_2, \dots, e_n)$
element construction	element $e_1$ { $e_2$ }
XPath steps	$e/\alpha::\nu$ (full axis feature)
⋮	⋮

- Expressions nest arbitrarily!

## A rather standard relational algebra suffices.

$\pi$	column projection, renaming
$\sigma$	row selection


## A rather standard relational algebra suffices.

$\pi$	column projection, renaming
$\sigma$	row selection
$\bowtie$	equi-join
$\times$	Cartesian product
$\dot{\cup}, \setminus$	disjoint union, difference
$\delta$	duplicate elimination

## A rather standard relational algebra suffices.


$\pi$	column projection, renaming
$\sigma$	row selection
$\bowtie$	equi-join
$\times$	Cartesian product
$\dot{\cup}, \setminus$	disjoint union, difference
$\delta$	duplicate elimination
$\rho$	row numbering

## A rather standard relational algebra suffices.

$\pi$	column projection, renaming
$\sigma$	row selection
$\bowtie$	equi-join
$\times$	Cartesian product
$\dot{\cup}, \setminus$	disjoint union, difference
$\delta$	duplicate elimination
$\varrho$	row numbering
	staircase join*
$\varepsilon, \tau$	element/text node construction*

\*Syntactic sugar; expressible by remaining operators.


## A rather standard relational algebra suffices.

$\pi$	column projection, renaming
$\sigma$	row selection
$\bowtie$	equi-join
$\times$	Cartesian product
$\dot{\cup}, \setminus$	disjoint union, difference
$\delta$	duplicate elimination
$\varrho$	row numbering
	staircase join*
$\varepsilon, \tau$	element/text node construction*
$\otimes$	arithm./comparison operator *

\*Syntactic sugar; expressible by remaining operators.



## A rather standard relational algebra suffices.


$\pi$	column projection, renaming
$\sigma$	row selection
$\bowtie$	equi-join
$\times$	Cartesian product
$\dot{\cup}, \setminus$	disjoint union, difference
$\delta$	duplicate elimination
$\varrho$	row numbering
	staircase join*
$\varepsilon, \tau$	element/text node construction*
$\otimes$	arithm./comparison operator *

- Operates on node (not tree!) level, 1NF relations.

\*Syntactic sugar; expressible by remaining operators.

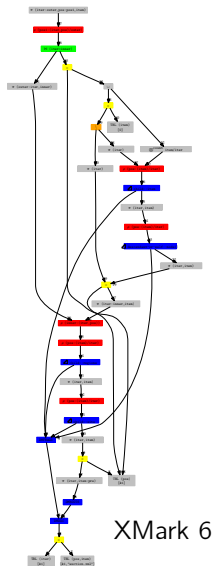
# Target Language: Relational Algebra

A rather standard relational algebra suffices.

$\pi$	column projection, renaming
$\sigma$	row selection
$\bowtie$	equi-join
$\times$	Cartesian product
$\dot{\cup}, \setminus$	disjoint union, difference
$\delta$	duplicate elimination
$\varrho$	row numbering
	staircase join*
$\varepsilon, \tau$	element/text node construction*
$\otimes$	arithm./comparison operator *

■ Operates on node (not tree!) level, 1NF relations.

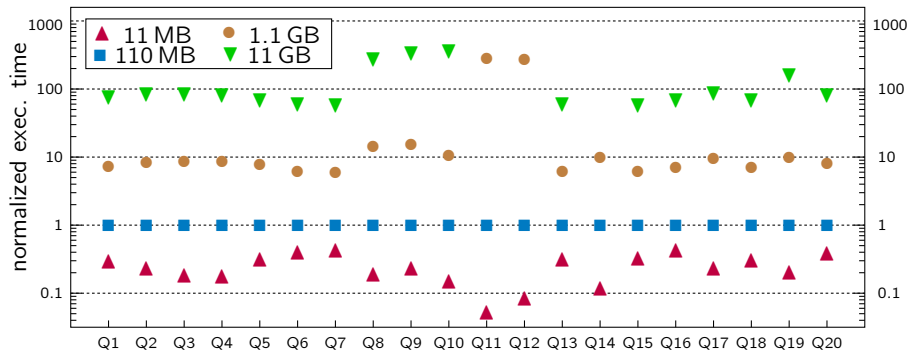
\*Syntactic sugar; expressible by remaining operators.



# Pathfinder Scalability

Version 0.8 of MonetDB/XQuery was released on May 30.

- Unsurpassed scalability, beyond 10 GB input document size.



The **Pathfinder project** is a joint effort of the **Technische Universität München**, **CWI Amsterdam**, and the **University of Twente**.

## Optimizations:

- Algebraic join detection
- Order awareness (avoid  $\rho$ )
- Use functional and multi-valued dependencies for algebraic optimization

## Open Source Implementation:

<http://www.pathfinder-xquery.org/>

- Backed by main memory DBMS **MonetDB**.



See you in **Demo Group 7** (today, 16:00; tomorrow, 16:00)!