

Towards micro-benchmarking XQuery

Ioana Manolescu
INRIA Futurs, France
ioana.Manolescu@inria.fr

Cédric Miachon
LRI - Université Paris-Sud 11,
France Cedric.Miachon@lri.fr

Philippe Michiels
University of Antwerp, Belgium
philippe.michiels@ua.ac.be

ABSTRACT

A substantial part of the database research field focusses on optimizing XQuery evaluation. However, optimization techniques are rarely validated by means of cross platform benchmarking. The reason for this is that there is a lack of tools that allows one to easily compare different implementations of isolated language features. This implies that there is no overview of which engines perform best at certain XQuery aspects, which in turn makes it hard to pick a reference platform for an objective comparison. This paper is a first step in a larger effort to bring an overview of the available implementations along with their strengths and weaknesses. It is meant to guide implementors in benchmarking and improving their products.

1. INTRODUCTION

In the recent past, a lot of energy has been spent on optimizing XML querying. This resulted in many implementations of the corresponding specifications, notably XQuery and XPath. Usually, little time and space is spent on thorough measurements across different implementations. This complicates the task of implementors to compare their implementations to the *state of the art* technology, since no one really knows what system actually represents it.

As is pointed out in [4], there are two possible approaches for comparing systems using benchmarks. Application benchmarks like XMark [15], XMach-1 [8], X007 [9] and XBench [17] are used to evaluate the overall performance of a database system by testing as many query language features as possible, using only a limited set of queries. As such, this kind of benchmarks are not very useful for XPath/XQuery implementors, since they are mainly interested in isolated aspects of an implementation that need improvement.

Micro-benchmarks, on the other hand, are designed to verify the performance of isolated features of a system. We believe that microbenchmarks are crucial in order to get a good understanding of an implementation. Moreover, it rarely happens that one platform is the fastest on all as-

pects. Only microbenchmarks can reveal which implementation performs best for isolated features.

Our focus is to benchmark a set of important XQuery constructs that form the foundation of the language and thus greatly impact the overall query engine performance. These features are:

- XPath navigation
- XPath predicates (including positional predicates)
- XQuery FLWORS
- XQuery Node Construction

The selected XQuery processors are chosen to represent both in-memory and disk-based implementations of the language as well as freelance engineering products.

We regard this paper as an initial step to a much larger micro-benchmarking effort that started with MemBeR [4] and which we hope to continue, using automated tools such as XCheck [5, 1]. This continuation involves the population of a repository with a large amount of ready-made micro-benchmarks as well as the benchmarking of many more platforms. We hope that this work can guide XQuery implementors to improve their products based on objective, thorough and relevant measurements.

Limitations This work does not attempt to make an absolute comparison of the tested XQuery implementations, nor does it mean to give an overview of the possible configurations of the implementations and their impact on the performance. Indeed, all tested systems have been run in their default configurations. Since not all implementations allowed an easy measurement of the the different query processing phases, we were forced to make an abstraction and focus on *query execution* and *result serialization* times. Hence, detailed discussions of *query compilation* and *tree construction* are out of this paper's scope. Further details on the methodology used for running the experiments can be found in Section 2.

2. SETTINGS

In this section, we present the documents (Section 2.1) and queries (Section 2.2 and 2.3) used for the performance measures in this paper, as well as the rationale for choosing them. Section 2.4 describes our hardware and software environment, and the system versions used.

All documents, queries, settings, and (links to) the systems used in these measures can be found at [2].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Proceedings of the First International Workshop on Performance and Evaluation of Data Management Systems (EXPDB 2006), June 30, 2006, Chicago, Illinois, USA
Copyright 2006 ACM 1-59593-463-4 ...\$5.00.

2.1 Documents

In order to have full control over the parameters characterizing our documents, we used synthetic ones, generated by the MemBeR project's XML document generator [4, 3]. MemBeR-generated documents consist of simple XML elements, whose element names and tree structure is controlled by the generator's user. Each element has a single attribute called `@id`, whose value is the element's positional order in the document. The elements have no text children.

Some of the systems we tested are based on a persistent store, while the others run completely in memory. While we are aware of the inherent limitations that an in-memory system encounters, we believe it is interesting to include both classes of systems in our comparison, since performant techniques have been developed independently on both sides, and the research community can learn interesting lessons from both. To enable uniform testing of all systems, we settled for moderate-sized documents of about **11 MB**, which most systems can handle well. As such, the stress testing of the systems below has a focus on query scalability, rather than data scalability.

Three documents have been used in these measures, whose structure is outlined in Figure 1. In this figure, white nodes represent elements, whose names range from `t1` to `t19`; black nodes represent `@id` attributes. All documents have the depth **19**, which we chose so that complex navigation can be studied, and in accordance with the average-to-high document depth recorded in a previous experimental study [14].

- The **exponential2.xml** document's size is **11.39 MB**. At level i (where the root is at level 1), the document has 2^{i-1} elements labeled `ti`.
- The **layered.xml** document's size is **12.33 MB**. The root is labeled `t1`, and it has **32768** children labeled `t2`. At any level i comprised between 3 and 19, there are **32768** nodes labeled `ti`. Each element labeled `ti`, with $3 \leq i \leq 18$, has exactly one child labeled `t(i + 1)`. Elements labeled `t19` are leaves.
- The **mixed.xml** document's size is **12.33 MB**. The root is labeled `t1`, and it has **3268** children labeled `t1`. Each such child (at level 2) has **10** children labeled (with equal probability) `t3`, `t4`, ..., `t12`. Nodes at levels comprised between 3 and 18 each have **1** child, labeled (with equal probability) `t3`, `t4`, ..., `t12`. At level 19, all nodes are leaves, and are labeled `t13`.

The rationale for choosing these documents is the following. The document **exponential2.xml** allows studying the impact of increasing number of nodes at a given level, on the performance of path traversal queries. At the same time, in this document, the size of a subtree rooted at level i is exponential in i . At another extreme, the document **layered.xml** has the same depth and approximate tree size as **exponential2.xml**, but the size of subtrees rooted at various levels depends only linearly on the level. The subtree shapes exhibited by both **exponential2.xml** and **layered.xml** are quite extreme; subtrees from real-life documents are likely to be somewhere in between. Controlling both the path depth and the size of the subtrees rooted at each depth is important, since these parameters have important, independent impacts on query performance: the first determines the performance of navigation queries, while the

second determines the performance of reconstructing (or retrieving) full document subtrees.

Our last document was chosen so as: (i) to be of overall size and aspects close to the two previous documents; (ii) to feature different tags uniformly distributed over many levels, thus allowing us to vary, in a controlled manner, the *structural selectivity* of various queries (by allowing some tags to range over increasingly large subsets of `{t1,t2,...,t13}`).

For simplicity, and unless otherwise specified, the documents were not characterized by type information.

2.2 XPath Queries

We measured on the document **exponential2.xml** seven parameterized XPath queries, denoted **Q1.1(n)**, **Q1.2(n)**, ..., **Q1.7(n)**, where $1 \leq n \leq 19$, as follows:

- **Q1.1(n)** is: `/t1/t2/.../tn`
This query retrieves nodes at increasing depths of the document. Its output size decreases as the roots of the returned subtrees move lower in the document. This query is designed to test the ability of the query processor to deal with increasing lengths of path expressions. For instance, intermediate materialization will have an increasing performance impact for longer queries. We also measured **Q1.1(n)** on **layered.xml**, which provided some interesting insights when compared to the results on the first document.
- **Q1.2(n)** is: `/t1/t2/.../tn/@id`
To distinguish the impact of navigation from the impact of subtree serialization in the output, we also use the query **Q1.2(n)**, which navigates at the same depth as **Q1.1(n)** but only returns simple attribute values.
- **Q1.3(n)** is: `(/t1/t2/.../tn)[1]/@id`
This query is used to see if query engines are able to take advantage of the `[1]` predicate to shortcut navigation as soon as a single node is found.
- **Q1.4(n)** is:
`(/t1/t2/.../tn)[position()=last()]/@id`
This query is similar to **Q1.3(n)**, but it uses the `[position()=last()]` predicate, which does not easily allow the same optimization as `[1]` if the engine is coded to navigate over the target nodes in the order dictated by XPath's semantics [10]. Measuring both **Q1.3(n)** and **Q1.4(n)** hints at the navigation optimization techniques supported in the engine.
- **Q1.5(n)** is: `/t1[t2/.../tn]/@id`
The queries **Q1.5(n)** aim at quantifying the impact of increasingly deeper navigation along existential branches. Once again this verifies whether query engines can get around materializing the predicate results and/or if they are capable to use shortcut evaluation, once a single predicate result has been found.
- **Q1.6(n)** is: `/t1[t2/.../tn]/t2/.../tn/@id`
Query **Q1.6(n)** presents an optimization opportunity (the existential branch can be suppressed without changing the query semantics); its results are to be interpreted together with those of **Q1.2(n)**
- **Q1.7(n)** is: `//tn`
Finally, query **Q1.7(n)** retrieves all elements of a given tag. Its results are to be compared with those of

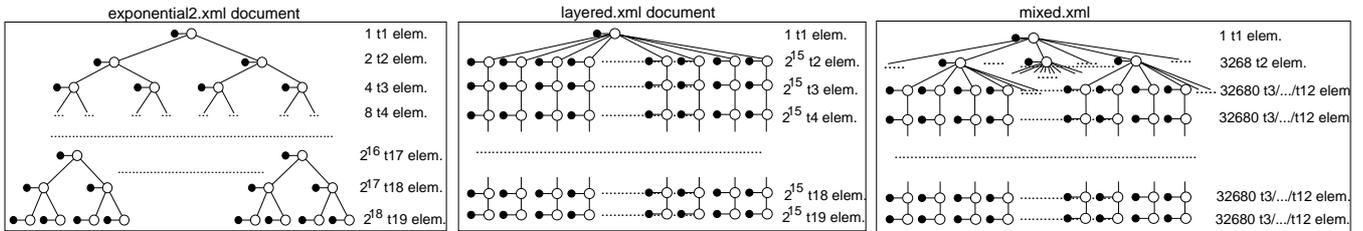


Figure 1: Outline of the documents `exponential2.xml`, `layered.xml` and `mixed.xml` used in the measures.

Q1.1(n), to see if the user’s knowledge of the depth of the desired elements simplifies the query processor’s task.

2.3 XQuery Queries

We measured on the document `mixed.xml` a set of six parameterized XQuery queries, denoted **Q2.1**(n), **Q2.2**(n), ..., **Q2.6**(n), where $0 \leq n \leq 9$, as follows:

- **Q2.1**(n) is:


```
for $x in /t1/t2
return <res>{ for $y in $x/**/.../*
return fn:data($y/@id) }
</res>
```

where the path expression to which $\$y$ is bound features n child navigation steps starting from $\$x$. These queries test the performance of increasingly deeper navigation *in the return clause*, while returning results of modest (and, on `mixed.xml`, constant) size.

- **Q2.2**(n) is:


```
for $x in /t1/t2
return <res>{ $x/**/.../* } </res>
```

where the path expression in the **return** clause features n child navigation steps starting from $\$x$. The queries **Q2.2**(n) combine increasingly deep navigation with decreasingly large subtrees to be copied in the output (recall that XQuery semantics [7] requires the content of newly created elements to be copied).

- **Q2.3**(n) is:


```
for $x in /t1/t2
return <res>{ $x//t3, $x//t4, ...,
$x//t(n+2) } </res>
```

The queries **Q2.3**(n) are designed to return results whose size is expected to increase with n , given that elements labeled **t3**, **t4**, ..., **t12** are uniformly distributed over levels 3-18 in `mixed.xml`. Moreover, the elements which **Q2.3**(n) must retrieve are scattered over the document. This allows us to verify how increasingly big and scattered results are dealt with by the processor.

- **Q2.4**(n) is:


```
for $x in /t1/t2
return <res>{ $x/*[position() ≤ n] } </res>
```

Q2.4(n) returns results whose size linearly grows with n , however in this case, the elements to be returned are grouped in contiguous subtrees in the original document. The performance of **Q2.3**(n) compared with that of **Q2.4**(n) provides interesting insight on the node clustering strategy used by the system (if any).

- **Q2.5**(n) is:


```
for $x in /t1/t2
return $x/*[position() ≤ n]
```

The queries **Q2.5**(n) are similar to **Q2.4**(n), however **Q2.5**(n) does not construct new elements. Strictly speaking, **Q2.5**(n) could have been expressed in XPath, but we keep it in the XQueries group for comparison with **Q2.4**(n). Given that **Q2.5**(n) does not construct new elements, there is an opportunity for a more efficient evaluation than in the case of **Q2.4**(n), since no tree copy operation is needed.

- **Q2.6**(n) have increasingly deeply nested **return** clauses. All the queries retrieve **t13** elements, and return them “wrapped” in increasingly deeper `<res>` elements. Rather than giving the general form, quite difficult to read, we provide here some examples:

Q2.6(0):

```
for $x in /t1/t2 return
<res>{for $x1 in $x/* return
<res>{ $x1//t13 } </res> }
</res>
```

Q2.6(1):

```
for $x in /t1/t2 return
<res>{for $x1 in $x/* return
<res>{for $x2 in $x1/* return
<res>{ $x2//t13 } </res> }
</res> }
</res>
```

Q2.6(2):

```
for $x in /t1/t2 return
<res>{for $x1 in $x/* return
<res>{for $x2 in $x1/* return
<res>{for $x3 in $x2/* return
<res>{ $x3//t13 } </res> }
</res> }
</res> }
</res>
```

Query **Q2.6**(n) returns subtrees consisting of $n + 2$ recursively nested `<res>` elements, each of which encloses some **t13** elements. on the document `mixed.xml`, will all return 3260 `<res>` elements, since there are 3260 **t2** elements in `mixed.xml`. Moreover, each such `<res>` elements will include copies of 10 **t13** elements. As n grows, however, the number of “layers” of `<res>` elements in which the **t13** elements are wrapped increases. The queries and the document have been chosen to capture the impact of result nesting only.

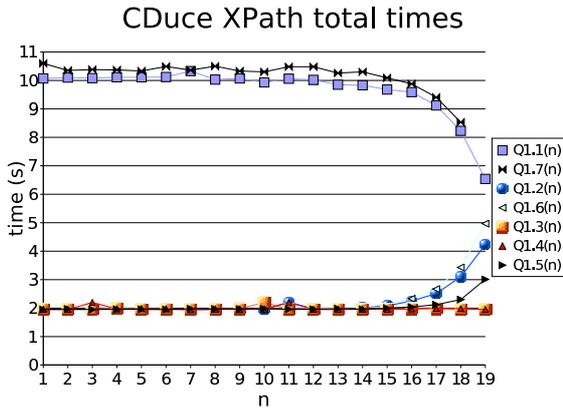


Figure 2: CDuce XPath total times on *exponential2.xml*.

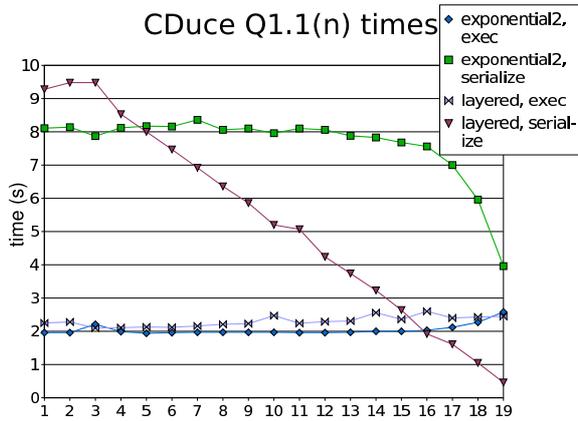


Figure 3: CDuce XPath component times for $Q1.1(n)$.

2.4 Hardware and software environment

Our measures were performed on four different computers. While ideally a single machine should have been used, the four computers have overall similar parameters. Also, in this paper we are mostly interested in identifying the tendency of each system's running time across increasingly complex queries, which tend to be quite stable, even if different computers are used. The machines are described by the following hardware and software parameters:

M1 is equipped with a 2.00 GHz Pentium 4, 512 MB of RAM, running Linux 2.6.12-10-386.

M2 is a DELL Precision M70 laptop, equipped with a 2.00 GHz Pentium, 1GB of RAM, running Linux 2.6.13.

M3 is equipped with a 3.00 GHz Pentium 4, 512 MB of RAM, running Linux Debian 2.6.16.

M4 is equipped with a 3.00 GHz Pentium 4 CPU, with 2026 MB of RAM, running Linux version 2.6.12.

We tested the following systems:

CDuce/CQL version 0.4.0. The system is implemented in OCaml. CDuce execution proceeds in two stages: first the query is compiled by invoking `cduce --compile`

CDuce XPath total times with precise DTD

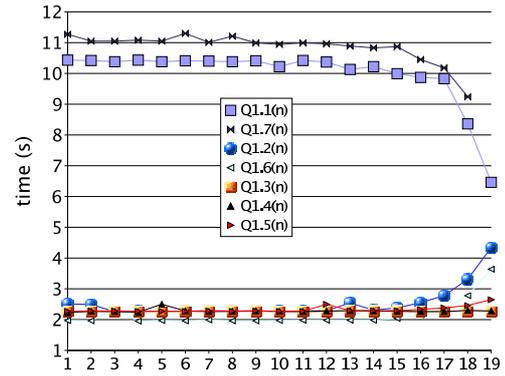


Figure 4: CDuce XPath total on *exponential2.xml* times in the presence of precise DTD information

CDuce XQuery times

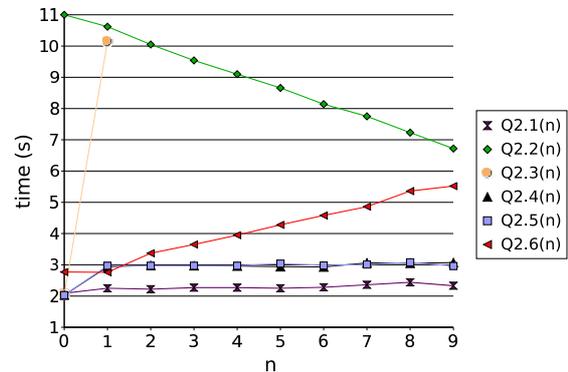


Figure 5: CDuce XQuery total times on *mixed.xml*.

`query.cd --obj-dir cdo`, then the query is executed by invoking `cduce --run query.cdo -I cdo`.

eXist version 1.0 beta 1. The system is implemented in Java. We first launched an eXist server using the `startup.sh` script part of the distribution. Then, we launched the execution of each query by running the `org.exist.examples.xmldb.XQueryExample` Java class, slightly modified to measure serialization time.

Galax version 0.5.0. The system is implemented in OCaml. The command line used to launch the experiment was: `galax-run -monitor-time on query.xq`.

MonetDB/XQuery version 0.10. The system is implemented in C++. We used a 32-bit compilation with no optimizations. A MonetDB server was launched using the `MServer` command, and queries were ran using the `MapiClient` utility.

QizX/open version 1.0. The system is implemented in Java. We launched the execution by calling `qizxopen_batch.sh query.xq > buf`, where `query.xq` is a file containing the query, and `buf` is a temporary buffer file receiving the output. The script `qizxopen_batch.sh` is part of the distribution.

Saxon version 8.6.1. The system is implemented in Java.

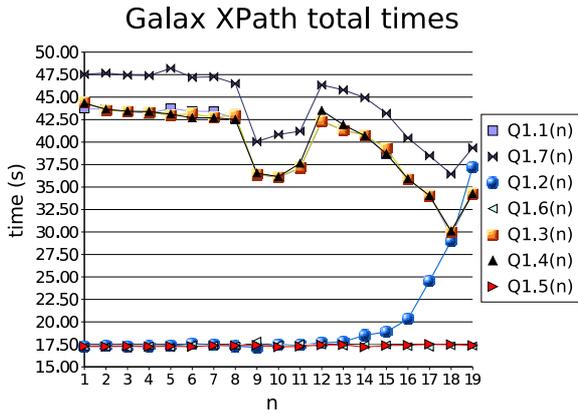


Figure 6: Galaxy XPath total times on *exponential2.xml*.

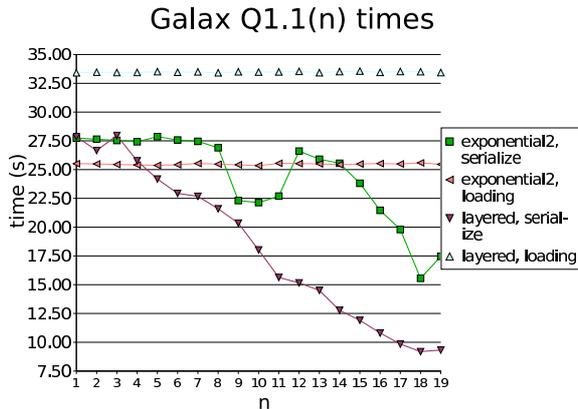


Figure 7: Galaxy XPath partial times on *exponential2.xml*.

The command line used was: `net.sf.saxon.Query -t -o result query >& times.`

We chose systems that (i) were freely available (if possible open source), (ii) had a user community and/or (iii) were the target of recent published research works. Our choice of systems includes some endowed with a persistent store (eXist and MonetDB), as well as purely in-memory systems (CDuce, Galax as used in these measures, QizX and Saxon). In this work, we did not specifically target our measures at disk-based retrieve times of disk-resident systems (although, of course, this aspect is interesting). Rather, we aimed at studying the performance of various algorithms implemented in the engines once they run in memory.

To that effect, we ran *each measure 4 times, and report the average of the last 3 (hot) runs.*

We are aware that more systems meeting our criteria exist, and plan to extend our tests to such systems in the near future.

3. RESULTS BY SYSTEM

This section discusses the results for each of the benchmarks per individual system.

3.1 CDuce

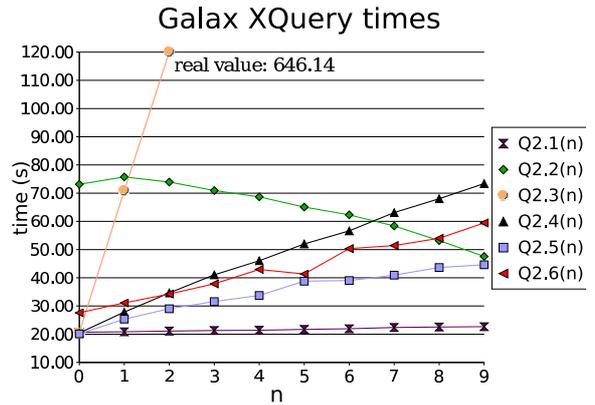


Figure 8: Galaxy XQuery total times on *mixed.xml*.

Figure 2 depicts the total running time of CDuce for our XPath queries on *exponential2.xml*. These measures were performed on **M3**.

CDuce (pronounce "seduce") is a general purpose typed functional programming language, specifically targeted to XML applications. It conforms to basic standards such as DTDs, Namespaces etc. Theoretical foundations of the CDuce's type system can be found in [12]. The system is implemented in OCaml.

Recently, a "Select-From-Where" syntax similar to XPath has been added for user convenience on top of CDuce [6]; it is translated into CDuce. The CDuce system, a sample demo, and online documentation can be found at [16]. An important characteristic of CDuce is its *pattern algebra*, which extends the XDuce [13] pattern algebra, allowing complex and powerful pattern matching. Pattern matching in CDuce is implemented by means of automata constructed "just-in-time".

CDuce has a strong type system, which enables static verification of safe program composition. Of course, CDuce can also evaluate queries in the absence of type information, using a more general automaton. A large subset of XQuery (excluding, for instance, function calls) is covered by CQL, and all the queries we measured in this paper are part of this subset. (Other XQuery features not considered here can be expressed directly in the CDuce language.) Among the XQuery features not supported is *node identity*: CDuce is value-based, that is, it does not distinguish two distinct nodes having the same serialized value (other than by their respective positions in the original document). In particular, there is no way to test if two variable bindings correspond to *the same* node, in XQuery sense [7].

The running times of **Q1.1(n)** and **Q1.7(n)** are comparatively more important, and decrease as n grows, due to the decreasing total size of the query result. Interestingly, **Q1.7(n)** causes a stack overflow for $n = 19$. This is because descendent navigation is implemented in CDuce via a stack that keeps all descendant matched on recursive function calls (and automata), unlike child, parent and sibling navigation, which are implemented by automata alone. However, the curves for **Q1.1(n)** and **Q1.7(n)** are roughly similar.

Both **Q1.3(n)** and **Q1.4(n)** have very short running time and are almost constant. The translation of these XPath queries to CQL, combined with the pattern matching concept underlying the system, allows to push the [1] and [posi-

`tion()=last()` predicates at all levels up, thus the efficient evaluation.

The time for **Q1.2**(n) exhibits an exponential growth as the number of returned nodes grows exponentially. The same observation holds for **Q1.6**(n), which is more expensive than **Q1.2**(n), although they are equivalent. The reason is that the pattern resulting from the translation of **Q1.6**(n) has twice the size of the pattern for **Q1.2**(n) (in other words, the pattern is not minimized to eliminate the redundant existential branch).

Q1.5(n) times tend to grow sensibly for large values of n . The reason is that the corresponding pattern's size grows linearly with n , and the curve reflects the complexity of automata-based evaluation for such patterns.

CDuce reported times are the sum of: **the automaton construction time**, determined by the query and the input document's DTD; **the execution time**, during which the results are constructed but not serialized yet; and **the serialization time**. We refine our analysis of CDuce running times in Figure 3, which plots execution and serialization time for **Q1.1**(n) on **exponential2.xml** and **layered.xml**. The serialization time linearly reflects the total result size, and the number of returned subtrees (the former dominates, however, in the case of **exponential2.xml**). The execution (retrieval) time is almost constant for increasing n , due to the efficient pattern matching operator.

In the above results (as in all other curves of this paper), no type information was present, and the automaton construction time was in all cases of about 20 ms.

Given CDuce's strong reliance on types (the evaluation is made by automata, which are built based on types), we re-ran the XPath queries, this time providing CDuce with a precise level-by-level DTD according to which **exponential2.xml** is valid. The results are shown in Figure 4. From **Q1.1**(n) to **Q1.7**(n), the running times are equivalent with the running times in Figure 2.

Figure 5 presents CDuce's running times on XQueries. **Q2.1**(n) grows very slowly with n , reflecting a moderately increasing navigation time, and a small serialization effort. **Q2.2**(n) running time decreases as the size of the returned subtrees decreases. **Q2.3**(n), which builds the larger results, failed to run for $n \geq 2$. The reason is that each serialized element in CDuce is written in an OCaml string variable, and the maximum size of such string variables is constrained by the language environment. Elements returned by **Q2.3**(n) grow larger with n , and outgrow at some point the available space. Clearly, this is an engineering problem, whose solution involves the usage of some buffered data structure continuously writing to a file.

Q2.4(n) and **Q2.5**(n) have roughly the same running time, since for CDuce, a newly constructed node is just another value (there is no need to deeply copy trees). Finally, **Q2.6**(n) grows linearly with the result nesting levels.

3.2 eXist

Figure 10 shows eXist's running times on our XPath queries. These measures were performed on **M2**.

The queries **Q1.1**(n) and **Q1.7**(n) have similar running times, although there is some difference for large n values. **Q1.2**(n) and **Q1.6**(n) are very close, reflecting the negligible influence of the extra existential branch in **Q1.6**(n). Both **Q1.3**(n) and **Q1.4**(n) are significantly faster than **Q1.2**(n), showing that eXist is capable of optimizing path

navigation based on global position predicates. **Q1.5**(n) is the fastest, yet it does feature an exponential growth with large n values.

All the curves tend to grow exponentially for large n values, whether the query returns exponentially many nodes (as **Q1.1**(n) and **Q1.2**(n) do, or a constant number of nodes (as **Q1.5**(n) does). A possible explanation for this is that navigation *traverses* nodes at all layers from 1 to n for each of the queries **Q1.1**(n)-**Q1.7**(n); the number of such nodes is exponential with n , and the cost of this traversal dominates XPath evaluation costs.

To verify this hypothesis, we take a closer look at the partial times recorded for **Q1.1**(n), on **exponential2.xml** and also on **layered.xml**. The *locate time* (reported by eXist as **execution**) and the *serialization time* (reported by eXist as **display**) are shown in Figure 9. For readability, we separated the locate times (quite small) from the serialization times.

Figure 9 shows that the locate time grows linearly on the **layered.xml** document, and exponentially with the **exponential2.xml** document. This is consistent with the hypothesis that eXist navigates top-down and thus traverses all nodes from the root to the targets of the XPath expression. Concerning serialization times, Figure 9 shows that it is directly correlated with the *serialized result size* (thus the decrease on the **layered** document), but also with the *number of returned subtrees* (thus the exponential increase on the **exponential2.xml** document).

Figure 11 shows eXist's running times on the XQueries we considered. For **Q2.3**(1) and **Q2.6**(0) we stopped the execution after failing to get any output in 180 seconds. Clearly, descendant navigation in return clauses is hard to handle for eXist.

For the remaining queries, we record almost constant times for **Q2.2**(n), **Q2.4**(n) and **Q2.5**(n).

For **Q2.2**(n), the execution time grows from 2 seconds ($n=0$) to 27 seconds ($n=9$) as longer paths are navigated, while the serialization time decreases correspondingly as smaller subtrees are output. The two tendencies almost compensate each other, thus the quasi-constant overall running time.

For **Q2.4**(n) and **Q2.5**(n), the execution and serialization times are both quite unchanged as n increases. This is a surprising result, for which we do not have a satisfactory explanation. We plan to approach the eXist developers for further explanations.

Finally, **Q2.1**(n) running times grow linearly with n . The execution time is the main factor here, ranging from 7.3s ($n=0$) to 43s ($n=9$); the serialization time is about unchanged, as same-volume results are returned.

From the detailed running times of **Q2.1**(n) and **Q2.2**(n), we may infer that *execution times* as reported by eXist also reflect navigation performed inside **return** clauses. "Execution time" as reported by some other systems may be more restricted (see Section 3.5). This highlights the importance of carefully measuring and interpreting various partial times reported by the engines.

3.3 Galax

These tests were run on machine **M1**. Galax is a main memory XQuery engine that implements the XQuery Data Model (XDM). XPath evaluation happens iteratively, i.e., using nested loops with intermediate sorting and duplicate elimination. Constructors are implemented as cursors and

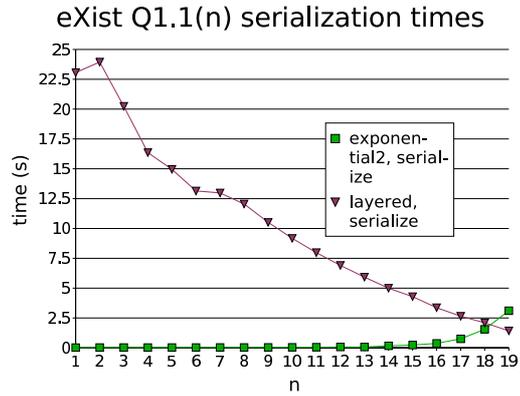
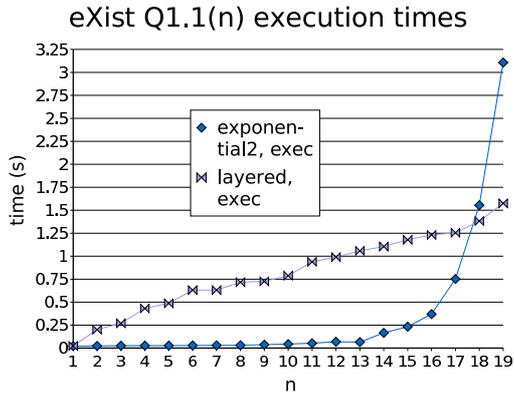


Figure 9: eXist XPath locate and serialization times on *exponential2.xml*.

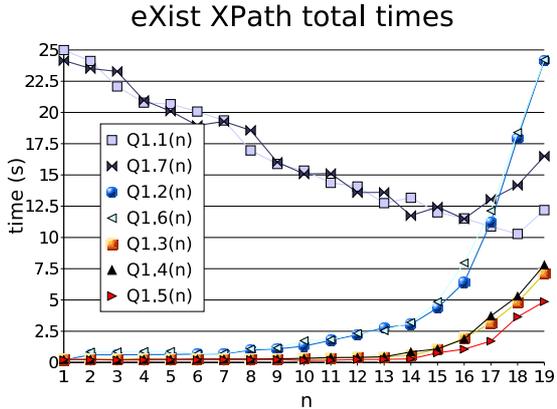


Figure 10: eXist XPath total times on *exponential2.xml*.



Figure 11: eXist XQuery times on *mixed.xml*.

should have a limited impact on query performance. The greatest drawback of Galax’s evaluation strategy is the need for materialization of intermediate results of path expressions, which requires a substantial amount of memory.

The results for the XPath queries are depicted in Figure 6. The result for **Q1.1(n)**, **Q1.3(n)**, **Q1.4(n)** and **Q1.7(n)** are quite similar. The curious fallback starting at $n = 9$ is believed to be an artifact of Galax serialization, which relies on OCaml pretty printing. This thesis is supported by more detailed monitoring information, given in Figure 7. The similarity of the results indicates that all four kinds of queries are handled in a similar fashion, implying materialization of all the result sequences. The exponential growth of query evaluation time is cancelled out by decreasing serialization overheads for the shrinking results. This effect is not noticeable for **Q1.3(n)** which selects only attribute nodes. The exponential growth here corresponds with the exponentially increasing result size. The constant behavior of queries **Q1.5(n)** and **Q1.6(n)** is an indication that Galax uses shortcut evaluation for predicates.

The XQuery results for Galax in Figure 8 show similar performance problems for **Q2.3(n)** as seen on other platforms, which are related to materializing large intermediate results in main memory. In the case of Galax, queries **Q2.3(n)** for $n \geq 2$ did not even finish within a reasonable time. Similar to MonetDB, Galax performs much better for

Q2.5(n), which returns similar result sizes but has no need for materializing the result in memory before serializing it. The other queries seem to scale along with the result size, since their performance is largely determined by the cost of serialization.

The execution times for Galax are relatively high for most queries. This is due to the fact that the Galax release used for benchmarking is quite old and due to a subtle bug related to node construction, which was revealed by these benchmarks.

3.4 MonetDB

Figure 12 plots MonetDB’s execution times on the XPath queries we considered. These measures were performed on **M4**.

Note that the times for **Q1.1(n)** and **Q1.7(n)** coincide, showing that descendent navigation does not pose particular difficulties for MonetDB. The times for **Q1.3(n)** and **Q1.5(n)** are very similar, and very low for all n values. The times for **Q1.2(n)** and **Q1.6(n)** are very low until $n = 16$, then an exponential growth projects the curves upwards (the actual value for $n = 19$ is shown on top of the graph). Similarly, **Q1.4(n)** exhibits an abrupt growth at $n = 11$.

The overall descending tendency of **Q1.1(n)** demonstrates that deep navigation does not pose difficulties for MonetDB; the dominating cost component here is clearly serialization, which decreases as n grows. Comparing **Q1.1(n)**

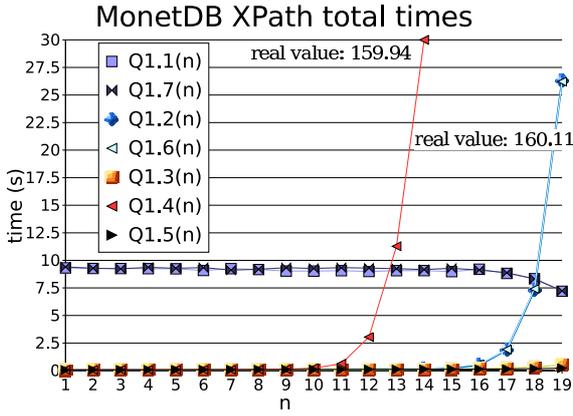


Figure 12: MonetDB XPath total times on *exponential2.xml*.

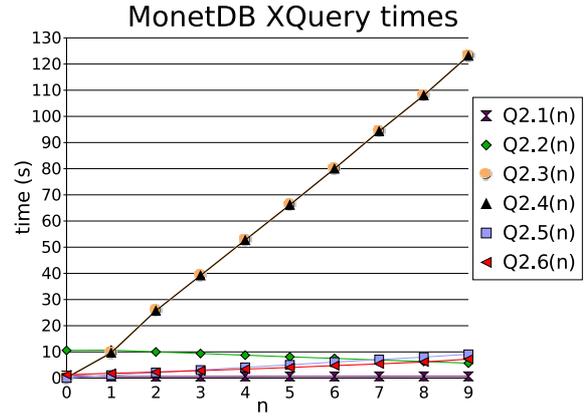


Figure 14: MonetDB XQuery total times on *mixed.xml*.

MonetDB Q1.1(n) times



Figure 13: MonetDB XPath component times on *exponential2.xml*.

and **Q1.2(n)** (on its initial segment) highlights the impact of the size of each returned subtree on the evaluation time; **Q1.2(n)** only returns simple data nodes, while **Q1.1(n)** returns more complex subtrees. The almost identical curves for **Q1.2(n)** and **Q1.6(n)** shows that the extra existential branch in **Q1.6(n)** is either eliminated at query compile time, or very efficiently executed.

The very low cost of **Q1.3(n)**, compared with the cost of **Q1.4(n)** which explodes above $n = 11$ and the cost of **Q1.2(n)** as a baseline, leads to the following observation. The `[1]` predicate after a long path expression in **Q1.3(n)** is exploited to make the evaluation more efficient than plainly enumerating all nodes of the path expression, and then returning just the first one. Thus, **Q1.3(n)** is stable and cheap, while **Q1.2(n)** (which *has* to enumerate exponentially many `@id` nodes) explodes at some point. On the contrary, the `[position()=last()]` predicate in **Q1.4(n)**, far from giving optimization opportunities, only makes matters worse than if no predicate is present, since **Q1.4(n)** costs rise before **Q1.2(n)** costs do.

The behavior exhibited by **Q1.2(n)**, **Q1.4(n)** and **Q1.6(n)** can be attributed to the handling of exponentially many nodes in intermediary XPath results. Curiously enough **Q1.2(n)** exhibits a significant performance difference with **Q1.1(n)**, indicating that the attribute axis is not handled

in the same way as the child axis¹.

Figure 13 plots MonetDB's detailed times for **Q1.1(n)**, on *exponential2.xml* and *layered.xml*. The partial times reflect, respectively, *query compile*, *query execution*, and *result serialization*. Figure 13 clearly shows that serialization is the single most important cost component for **Q1.1(n)**; both query compilation and query execution (understood here as the time to locate the return nodes only) are very small (more precisely, smaller than 0.2s across all documents and n).

MonetDB execution time on XQueries is depicted in Figure 14. **Q2.3(n)** and **Q2.4(n)** are of similar costs, and by far the most expensive. Both these queries return large results, constructed either from contiguous subtrees (**Q2.4(n)**) or from elements scattered all over the document (**Q2.3(n)**). Clearly, the location of the elements to be copied in the output does not have a big impact on evaluation performance. In contrast, **Q2.5(n)** which returns similar data volumes takes much less time. This shows that MonetDB correctly realizes no tree copy is needed for this query, but only a serialization of the *existing* returned nodes. The time of **Q2.5(n)** is of the same order of magnitude as the times for **Q1.1(n)** (Figure 12). Thus, we conclude that tree copying is significantly more expensive than serializing existing nodes in MonetDB (while keeping a linear scale-up), and that MonetDB correctly recognizes the situations where each of these should be used.

3.5 QizX

Figure 15 shows QizX execution times for the XPath queries we considered, on the *exponential2.xml* document. All measures of this section were performed on **M2**.

For low n values, the execution times of **Q1.1(n)** and **Q1.7(n)** decrease with n , due to the decrease in the volume of returned data. **Q1.1(n)** and **Q1.7(n)** exhibit an exponential growth with large n values, justified by the exponential increase in the number of returned nodes. Interestingly, the times for the `//ti` queries are quite different from the times of the equivalent query based on child steps, suggesting that QizX's search for elements to return at any depth in the document actually does search at all levels (no

¹The explosion of both **Q1.2(n)** and **Q1.4(n)** is believed to be due to a flaw in attribute handling and is currently under investigation by the MonetDB developers.

QizX XPath total times

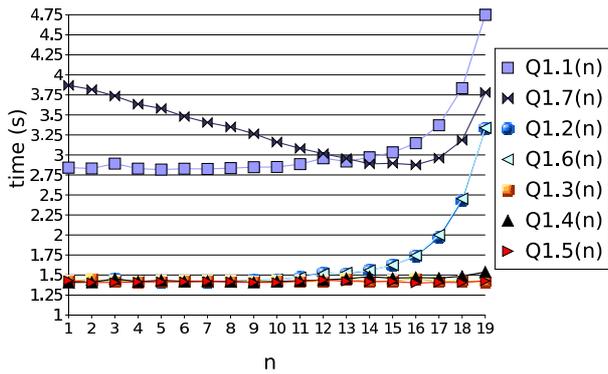


Figure 15: QizX XPath total times on *exponential2.xml*.

QizX Q1.1(n) times

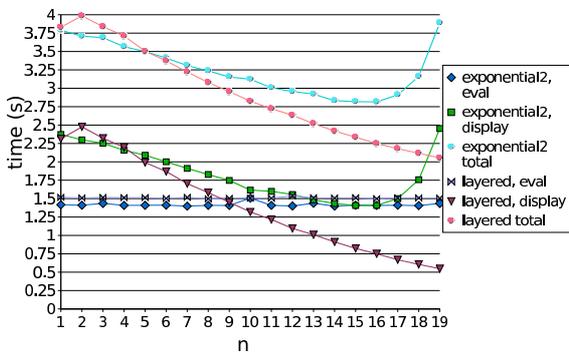


Figure 16: QizX XPath component times on *exponential2.xml*.

structural or tag-based index seems to be in place).

QizX allows to isolate in a measure:

the evaluation time: presumably the time spent processing the navigation (XPath) part of a query (in other words, the locate time);

the display time: the time needed to produce a serialization of the returned nodes (in other words, the serialization time).

We refine our understanding of QizX's behavior on **Q1.1(n)** by measuring the evaluation and the display times both on **exponential2.xml** and on **layered.xml**. Figure 16 shows the results. We notice that QizX's reported evaluation time is almost unchanged for all documents and path depths, which is quite surprising, and leads to believe that it corresponds to some document pre-processing time (e.g. to build a tree representation of the document and/or index all nodes by their paths etc.). The display (probably serialization) time decreases almost linearly with the growth of n , which is the expected behavior for the **layered.xml** document. However, the display time on **exponential2.xml** shows an exponential increase for large n values. On this document, the size of the query result decreases as n grows, and so does the number of nodes *included in the result subtrees*, however

QizX XQuery times

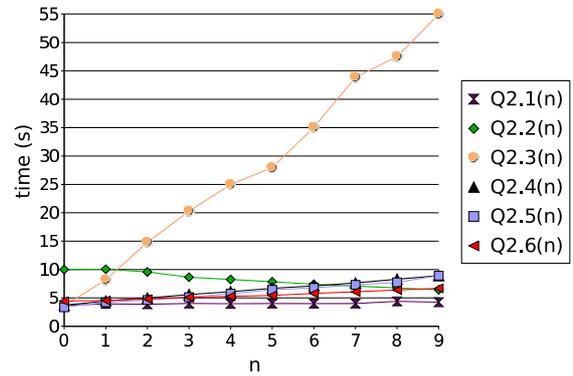


Figure 17: QizX XQuery total times on *mixed.xml*.

QizX XQuery times

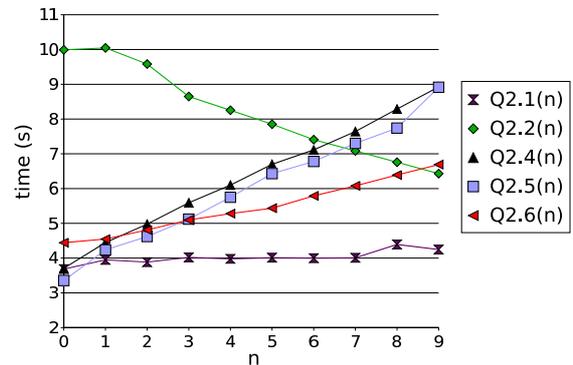


Figure 18: Same as Figure 17, excluding **Q2.3(n)**.

the *number of result subtrees* increases exponentially with n . Thus, the serialization complexity of QizX is strongly influenced *number of result subtrees*, and also (to a lesser extent) by the *number of nodes in the result subtrees*.

Figure 17 shows QizX's total XQuery execution times on **mixed.xml**. For all queries, QizX's reported evaluation time was between 12ms and 30ms, thus the times in Figure 17 practically coincide with its display time. Such insignificant (and constant) evaluation times lead to believe that for an XQuery, QizX reports as evaluation time the time needed to process *the for clause of the outermost FLWOR expression* only, while the rest of the query execution is reported as display time.

Figure 17 shows that **Q2.3(n)**, which returns massive results scattered over the whole document, is the single most expensive query. Comparing this with the times for **Q2.4(n)**, which returns similar data volumes but contiguous subtrees, allows to infer that `//` navigation, in the **return** clause as well as in **for** clauses (or XPath) brings quite a penalty to QizX.

Figure 18 plots the same times for all XQueries except **Q2.3(n)**. **Q1.1(n)** takes almost constant time; given that the cardinality and size of its result is constant as n grows, we can conclude that deep step-by-step navigation has no big impact on QizX performance, and confirm that the size and cardinality of the result strongly impacts serialization. The decreasing times of **Q2.2(n)**, producing shallower trees as n

grows, support the same assertion. $Q2.4(n)$ and $Q2.5(n)$ are quite close, with a small (but persistent) overhead of $Q2.4$ over $Q2.5$, probably due to the construction of a new `<res>` node as required by $Q2.4$. However, the difference seems too small to reflect fully copying the subtrees included in `<res>`.

The overhead of nested `<res>` elements in $Q2.6(n)$ is noticeable, while the scale-up remains linear.

3.6 Saxon

Figure 19 shows Saxon’s running time on our XPath queries. These measures were performed on **M4**.

A first remark is that $Q1.1(n)$ and $Q1.7(n)$ coincide, demonstrating that descendent navigation is quite efficient in Saxon; these results indicate the presence of a path index and

pipelined evaluation. Essentially this is done with a straightforward “nested loop” or rather, a Jackson-inverted nested loop using a cascading set of iterators. The main trick that Saxon uses is efficient testing of the element names by comparing integer fingerprints. Also, a lot of effort goes into static analysis of the expression to decide whether a sort and duplicate elimination is needed or not.

Compare queries $Q1.2(n)$, returning all `@id` attributes at a given depth, $Q1.3(n)$, returning just the first `@id` attribute, and $Q1.4(n)$, returning just the last `@id` attribute at a given depth. $Q1.2(n)$ is the most expensive. $Q1.4(n)$ is faster, but still exhibits an exponential growth with large n values, which allows to infer that all `@id` nodes at a given level are enumerated, prior to choosing the last one. Finally, $Q1.3(n)$ is the fastest, and runs in almost constant time as n grows. This hints to a short-circuited evaluation of path queries followed by [1], which stops after the first node at each level is found.

$Q1.5(n)$ is fast and runs almost in constant time, showing the existential predicate of $Q1.5$ is efficiently handled. Finally, $Q1.6(n)$ has running times very similar to those of $Q1.2(n)$, confirming the efficient handling of the existential branch.

Saxon’s times on XQuery queries are reported in Figure 20. The standalone, expensive query is $Q2.3(n)$, which copies deep trees, starting at locations scattered over the document. In contrast, $Q2.4(n)$, whose copied subtrees are smaller and are found at contiguous locations, is much faster, and $Q2.5$, which does not need to copy subtrees, is yet a little bit faster.

4. CONCLUDING REMARKS

In this section, we summarize some of our experiments’ conclusions, and hint to avenues for future work.

4.1 Methodology Lessons Learned

A first observation is that benchmarking needs to be performed for a quite substantial number of data points. Failing to do so comes at the risk of not picking up important facts. For instance, in the QizX case exponential behavior of XPath evaluation only becomes apparent for path expressions that are long enough (see Figure 16).

Another important part of our approach is to vary only one benchmark parameter at a time. For instance, varying both document size and query size at the same time may cause effects that cancel each other out, eventually blurring or hiding the impact of the separate changes.

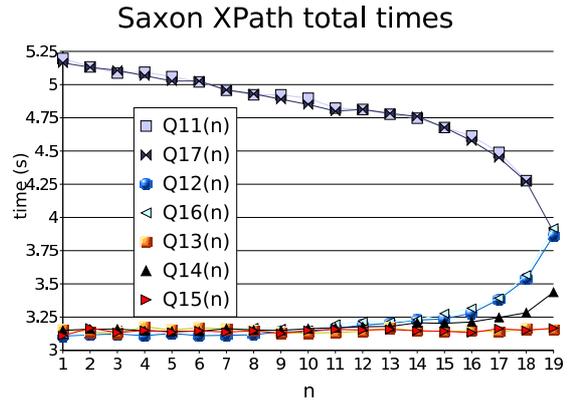


Figure 19: Saxon XPath total times on *exponential.xml*.

We also believe it is important to decompose query evaluation times into their components, as this is the only way to understand the impact of several parameters on the times. For instance, for Galax (Figure 7) and QizX (Figure 16) we have seen that query evaluation times are sometimes dwarfed by the time needed to serialize the result. It would be quite inappropriate to only report query evaluation time for such queries and systems, ignoring serialization. It would just be puzzling to report the overall time, without checking the respective evolution of its components.

We notice, however, that current system releases make it difficult or impossible to properly evaluate component times. The names and interpretations of such components time also vary. For instance, while the separation between “execution” and “serialization” seemed clean in all systems (and is reliable when measuring XPath queries), in the case of QizX, this separation seems arbitrary in the case of XQueries (recall Section 3.5). Extreme caution is therefore advised to the careful tester; few assumptions should be made on what a system does, or what its intermediate times mean, as long as these assumptions have not been checked with the system developers !

4.2 Performance Lessons Learned

Our study does not warrant a claim of having found *the best XQuery processor* among the systems we tested. It seems likely that different systems perform well at different features and under different circumstances. Compare for instance, MonetDB (Figure 12), performing quite badly on $Q1.4(n)$, with QizX (Figure 15) which is clearly faster, while on query $Q2.2(n)$ MonetDB (Figure 14) and QizX (Figure 17) have very similar performance.

Tradeoffs between features and efficiency are sometimes visible. For instance, CDuce sacrifices node identity for efficiency.

The impact of the implementation language on performance is quite limited. We might have expected a system implemented in C++ to perform some order of magnitude faster, however Java and OCaml times are very competitive and sometimes even better. (While our measures were performed on different machines, their computing power was quite uniform, which we believe enables such rough comparisons). This is reason for optimism, as we would indeed prefer algorithmics and efficient techniques to have a big-

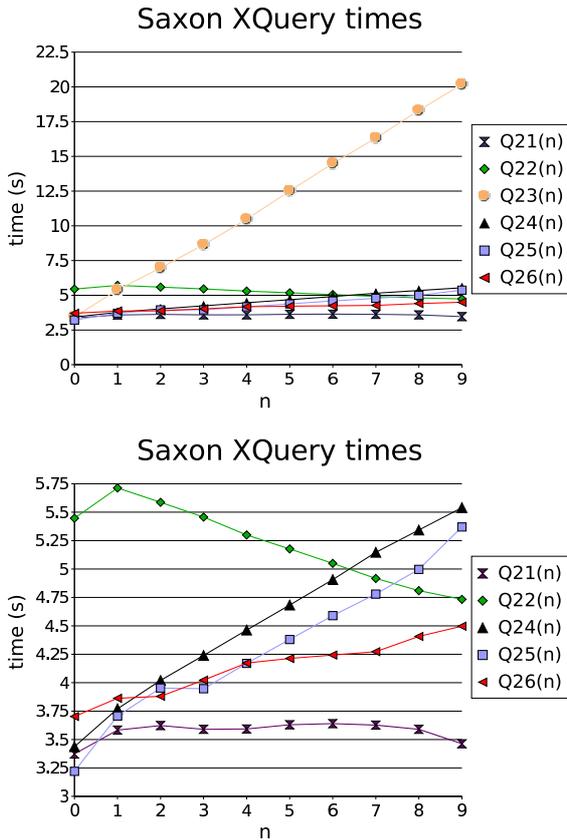


Figure 20: Saxon XQuery total times on *mixed.xml*. The second graph excludes **Q2.3(n)**.

ger impact on performance than the simple choice of the development language.

The performance of *child axis navigation*, a basic XPath feature, can be assessed by inspecting the execution time of **Q1.1(n)** (not including serialization), as shown in Figures 3, 9, 7, 13, and 16). Consider the behavior of **Q1.1(n)** on *exponential2.xml* and *layered.xml*, respectively:

- On *exponential2.xml*, the number of returned nodes is in $O(2^n)$, and the number of nodes visited by a naïve XPath evaluation strategy is also in $O(2^n)$.
- On *layered.xml*, the number of returned nodes is constant (and equal to the number of nodes returned by **Q1.1(14)**), whereas the number of nodes visited by a naïve XPath evaluation strategy is in $O(n)$.

A linearly increasing execution time for **Q1.1(n)** on *layered.xml* may show that the engine indeed implements a naïve top-down XPath evaluation strategy, which visits all intermediate nodes between the root and the target nodes. This is the case for eXist (Figure 9). On the contrary, for systems such as MonetDB (Figure 13), we see that the execution time is extremely small (and roughly constant). Other systems, such as QizX (Figure 16) and CDuce (Figure 3), exhibit constant time for finding the target nodes on *layered.xml*. This hints to the usage of efficient techniques for locating the nodes part of the query result. Observe that the retrieval time on **Q1.1(n)** needs to exhibit an exponential growth due to the number of returned nodes, but for sys-

tems such as MonetDB and QizX, the growth is negligible (unlike for eXist).

The impact of imprecise navigation (expressed with `// steps`) can be assessed by comparing **Q1.1(n)** with **Q1.7(n)**. In our tests, these queries are equivalent, while the second uses `//`. There is no noticeable difference among these queries for MonetDB and Saxon, which shows that these systems do not need to visit irrelevant nodes when evaluating **Q1.7(n)**. The other four systems exhibit some performance differences, most noticeable in the case of QizX.

Attribute navigation is well supported by all systems as can be seen by inspecting **Q1.2(n)** and **Q1.6(n)**. The exception is a performance problem of MonetDB, which was being solved at the time of this writing.

Early-stop optimizations allow the evaluation of queries like **Q1.3(n)** to proceed much faster than **Q1.2(n)**. This is the case of all systems except Galax, hinting to the fact that it needs to materialize all intermediate result sequences.

Existential branches or path predicates exhibit different performance depending on the tested system. This can be assessed by comparing **Q1.2(n)** with **Q1.6(n)**, as well as comparing **Q1.2(1)** with **Q1.5(n)**. Each of these pairs contains equivalent queries on the tested documents. These times coincide for eXist (Figure 10), MonetDB (Figure 12), QizX (Figure 15), and Saxon (Figure 19). However, **Q1.6(n)**, featuring an existential test, is more expensive for CDuce and Galax, the difference for Galax being quite important for large n values. In the case of CDuce, the bulk of the work in answering an XPath query is spent matching a corresponding tree pattern. This implies that all pattern embeddings in the data tree have to be found, whether or not they are actually returned.

Result serialization times corresponding to our queries are an important component of total running times, and in some cases, the dominant one (see, for instance Figure 13). This depends of course on the chosen queries and data, however our tests stress the importance this (often ignored) part of query execution times may take. Overall, serialization time reflects the number and size of the returned subtrees. The impact of some OCaml (or system programming) issues leads either to surprising performance variations (Figure 7), or to unfeasible queries such as **Q2.3(n)** for CDuce (Figure 5), eXist (Figure 11) and Galax (Figure 8).

Complex result construction such as required by **Q2.3(n)** and **Q2.6(n)** is related to the problem of serializing large results; **Q2.6(n)** constructs more complex new trees, but of smaller size. Interestingly, CDuce (Figure 5) and Galax (Figure 8) handle **Q2.6(n)** well, whereas eXist doesn't (Figure 11).

The above two observations highlight the work that some systems still need to scale up their serialization. Clearly, efficient algorithms exist, since MonetDB (Figure 14), QizX (Figure 17) and Saxon (Figure 20) handle such queries well.

New node creation is also an interesting feature in our measures. Queries **Q2.4(n)** and **Q2.5(n)** differ only by the fact that **Q2.4** constructs new trees (which requires copying some `t2` descendants) whereas **Q2.5** returns these nodes from the original documents. Three strategies are possible:

1. The system copies the trees for both queries. This would be incorrect for **Q2.5**, however if the query result is used directly serialized (as opposed to a "live" XQuery Data Model instance [11]), the incorrectness will go unnoticed. This is the strategy employed by

CDuce (Figure 5), which does not actually implement node identity.

2. The system copies the tree for **Q2.4** and does not copy them for **Q2.5**. This seems to be the strategy of eXist (Figure 11), Galax (Figure 8), MonetDB (Figure 14), QizX (Figure 18) and Saxon (Figure 20).
3. The system does not copy the trees neither for **Q2.4** nor for **Q2.5**. This is incorrect for **Q2.4**, however it will go unnoticed if query results are used directly serialized. None of the tested systems seems to use this strategy – although, in our setting, it would be optimal.

We also remark the the good performance of main memory implementations, notably Saxon, for the moderate-sized documents used in the experiments.

4.3 Future Work

As mentioned earlier, this paper is only a first in a greater effort, which targets a better understanding of the drawbacks and advantages of XPath/XQuery implementation strategies. An important step towards such a better understanding is the development of many more microbenchmarks, to reveal detailed information regarding the performance of implementations at isolated language features. We need to test more systems, such as e.g. BerkeleyDB/XML and Timber, and encourage developers to use our microbenchmarks to identify potential bottlenecks in their implementations. The “bad” cases for Galax and CDuce detected in this paper were news to their implementors. In the Galax case, these tests have revealed a subtle bug, which caused too much memory to be allocated, which in turn was also compromising time performance.

A platform for easy comparison of XQuery engines is currently under development. One thing that we can take away from the results of our tests is that a more clear separation of query evaluation times and serialization times is desirable. This could be included in the platform mentioned above by doing an additional run without serialization. Unfortunately, just like monitoring serialization separately, such is not supported by all engines. The same holds for document loading times in the case of main memory engines, which can be easily measured by just loading the document.

Important additional tests to be performed are system stress tests, that identify the operational boundaries of XPath and XQuery implementations. Other interesting aspects to be tested include: the impact of XML Schema information on navigation queries; handling of atomic values, value joins etc.; handling of increasing branching factors by XPath etc.

Acknowledgments. Philippe Michiels is supported by the IWT (Institute for the Encouragement of Innovation by Science and Technology) Flanders, grant number 31016. Ioana Manolescu is partly supported by a French Government “Young Researcher” grant.

We would like to thank Loredana Afanasiev for helping out on running some of the benchmarks, as well as Giuseppe Castagna, Véronique Benzaken and Alain Frisch for their help in interpreting CDuce results. We also thank Michael Kay for clarifying some of the Saxon results and Peter Boncz for his feedback on the MonetDB numbers.

5. REFERENCES

- [1] L. Afanasiev, M. Franceschet, M. Marx, and E. Zimuel. Xcheck: A platform for benchmarking XQuery engines (demo), 2006. To appear.
- [2] L. Afanasiev, I. Manolescu, C. Miachon, and P. Michiels. Micro-benchmarking XQuery experiments page, 2006. <http://www-rocq.inria.fr/~manolescu/microbenchmarks.html>.
- [3] L. Afanasiev, I. Manolescu, and P. Michiels. Member: A micro-benchmark repository, 2005. <http://ilps.science.uva.nl/Resources/MemBeR/>.
- [4] L. Afanasiev, I. Manolescu, and P. Michiels. Member: A micro-benchmark repository for XQuery. In *XSym*, volume 3671 of *Lecture Notes in Computer Science*, pages 144–161. Springer, 2005.
- [5] L. Afanasiev and M. Marx. XCheck, an automated XQuery benchmark tool, 2005. <http://ilps.science.uva.nl/Resources/XCheck>.
- [6] V. Benzaken, G. Castagna, and C. Miachon. A full pattern-based paradigm for XML query processing. In *PADL 05, 7th International Symposium on Practical Aspects of Declarative Languages*, number 3350 in LNCS, pages 235–252. Springer, Jan. 2005.
- [7] S. Boag, D. Chamberlin, M. Fernández, D. Florescu, J. Robie, and J. Siméon. XQuery 1.0 An XML Query Language, W3C Working Draft, April 2005. <http://www.w3.org/TR/xquery>.
- [8] T. Böhme and E. Rahm. Xmach-1: A benchmark for XML data management. In *Proceedings of BTW2001, Oldenburg, 7.-9. Mz, Springer, Berlin*, March 2001.
- [9] S. Bressan, G. Dobbie, Z. Lacroix, M. Lee, Y. Li, U. Nambiar, and B. Wadhwa. X007: Applying 007 benchmark to XML query processing tool. In *CIKM*, pages 167–174. ACM, 2001.
- [10] W. W. W. Consortium. XML path language (XPath) version 2.0 – W3C Working Draft, 2005. <http://www.w3.org/TR/xpath20/>.
- [11] W. W. W. Consortium. XQuery 1.0 and XPath 2.0 Data Model (XDM), 2006. <http://www.w3.org/TR/xpath-datamodel.html>.
- [12] A. Frisch, G. Castagna, and V. Benzaken. Semantic Subtyping. In *Proceedings, Seventeenth Annual IEEE Symposium on Logic in Computer Science*, pages 137–146. IEEE Computer Society Press, 2002.
- [13] H. Hosoya and B. Pierce. XDuce: A typed XML processing language. *ACM Transactions on Internet Technology*, 3(2):117–148, 2003.
- [14] L. Mignet, D. Barbosa, and P. Veltri. The XML web: A first study. In *WWW Conference*, 2003.
- [15] A. R. Schmidt, F. Waas, M. L. Kersten, D. Florescu, M. J. Carey, I. Manolescu, and R. Busse. Why and How to Benchmark XML Databases. *SIGMOD Record*, 3(30):27–32, 2001.
- [16] The CDuce project web site. <http://www.cduce.org>.
- [17] B. Yao, T. Özsu, and N. Khandelwal. Xbench benchmark and performance testing of XML DBMSs. In *ICDE*, pages 621–633. IEEE Computer Society, 2004.