

## Saxon - Support #5124

### Unexpected error "Cannot read a document that was written during the same transformation" - race condition

2021-10-13 10:11 - Johan Gheys

<b>Status:</b> New	<b>Start date:</b> 2021-10-13
<b>Priority:</b> Low	<b>Due date:</b>
<b>Assignee:</b>	<b>% Done:</b> 0%
<b>Category:</b>	<b>Estimated time:</b> 0:00 hour
<b>Sprint/Milestone:</b>	<b>Spent time:</b> 0:00 hour
<b>Legacy ID:</b>	<b>Fixed in Maintenance Release:</b>
<b>Applies to branch:</b>	<b>Platforms:</b>
<b>Fix Committed on Branch:</b>	
<b>Description</b>	
We use Saxon-EE 10.5 and one of our transformations is giving for the first time the unexpected error "Cannot read a document that was written during the same transformation" (see log file error-2021-10-12.txt). A simple restart fixed the problem. Probably seems closely related to issue 4934?	
<b>Related issues:</b>	
Related to Saxon - Bug #5130: `saxon:discard-document()` should be synchronized	<b>Resolved</b> <b>2021-10-15</b>

#### History

##### #1 - 2021-10-15 11:14 - Michael Kay

The first sign that things are wrong here is the message

```
Some child xsl:result-document threads have not finished
```

which should never happen.

Before that there are three messages of the form

```
*** MISSING file:/var/opt/app/arte2/u00p/interfac/exporter/publish/20211012-1726/formation/88_movement_HE_527.xml
```

which might well provide a clue.

The message

```
XTRE1500 Cannot read a document that was written during the same transformation:
  file:/var/opt/app/arte2/u00p/interfac/exporter/publish/20211012-1726/formation/88_movement_HE_527.xml
```

is actually a little inaccurate. It happens on a doc() call when the URI passed to doc() is unavailable, either because it was previously written using xsl:result-document, or for a variety of other reasons. The immediate cause is that the URI is present in a controller-owned collection called (misleadingly) allOutputDestinations, which is updated not only by xsl:result-document but also by fn:doc() and saxon:discard-document(). I suspect that if two threads both call doc() and subsequently call saxon:discard-document() on the same URI then it might be possible for the URI to be left in this collection at the end.

I'll need to think about that possibility a bit more carefully. If I'm right, then I think the only solution might be to introduce a function (say saxon:doc(uri, map{'stable':false})) that's equivalent to doing doc() followed by discard-document(), but does it as an atomic action.

In fact I see that the implementation of saxon:discard-document() isn't even synchronised within itself, which it certainly should be. That's easy to fix, but whether it's enough to solve the problem I don't know.

None of this explains the disquieting message Some child xsl:result-document threads have not finished. Is that something you see elsewhere in your log files?

##### #2 - 2021-10-15 12:55 - Michael Kay

I raised bug [#5130](#) regarding synchronization of saxon:discard-document(). It's quite hard to know whether that's actually a contributor to the current problem, but it's worth fixing anyway.

##### #3 - 2021-10-15 12:56 - Michael Kay

- Related to Bug #5130: `saxon:discard-document()` should be synchronized added

#### #4 - 2021-10-15 15:21 - Johan Gheys

Hi Michael, thanks for the investigation. Every \$runs/run is creating an xsl:result-document (see line 27 of select-trains-and-movements-(proxy).xslt) and I thought that error message was the result of an earlier problem (probably the reading problem caused by saxon:discard-document()) like you suggests).

#### #5 - 2021-11-07 00:27 - Michael Kay

Noted also that allOutputDestinations is a (thread-unsafe) HashSet which is updated by various non-synchronised methods such as XsltController.addAvailableOutputDestination()andXsltController.removeAvailableOutputDestination(). We need to use a thread-safe set/map here. But again, I'm not sure that's enough: the possible interactions between doc-available(), doc(), and discard-document() are quite complex.

#### Files

---

error-2021-10-12.txt	36.4 KB	2021-10-13	Johan Gheys
shared-functions.xslt	66.2 KB	2021-10-13	Johan Gheys
select-trains-and-movements-(proxy).xslt	15.3 KB	2021-10-13	Johan Gheys