

Apache™ FOP Design: Optimisations

Version 1298724

by Keiron Liddle

Table of contents

1 Introduction.....	2
---------------------	---

1 Introduction

Apache™ FOP should be able to handle very large documents. A document can be supplied using SAX and the information should be passed entirely through the system, from fo elements to rendered output as soon as possible.

A top level block area, immediately below the flow, can be added to the page layout as soon as the element is complete.

The fo elements used to construct a page can be discarded as soon as the layout for the page is complete. Some information may be stored in the area tree of the page in order to handle unresolved page references and links.

Once the layout of a page has been completed, all elements are fully resolved, then the page can be rendered. Some renderers may support out of order rendering of pages.

The main problem that will remain is that any page with forward references will need to be stored until the reference is resolved. This means that the information contained in the page should be as minimal as possible.

Line areas can be optimised once the layout for the line has been finalised. Consecutive characters with the same properties can be combined into a "word" to hold the information with limited overhead.

If there are a large number of pages where forward references cannot be resolved then a method of writing a page onto disk could be used to save memory. The easiest way to achieve this is to make the page and all children serializable.