White Paper

Batch vs. Online PDF Conversion Guidelines

We often have customers ask us whether they should convert their print files to PDF in a batch doing their print cycle, or whether they should store the print data in its original format and do the conversions online.

There are a number of considerations which should be taken into account when planning on how any given customer statement application should be managed. Many of these involve tradeoffs which must be considered in terms of document look and feel versus speed of delivery and storage space required. Based on the above statistics and our experience with applications in customer environments, the following are some guidelines which can be used for planning purposes.

The first step is to gather information and statistics about the statement application. The following information should be gathered, or estimated. Size of Metacode file. The size of the statements when they are split into individual metacode files will vary by application, based on density of text and the presence of page interleaved graphics. You can run the splitter and determine the average statement file size.

Size of Converted PDF file

After the statements are converted into PDF files, the PDF file may contain a number of objects which affect the file size. The primary objects which increase file size are:

- The size and number of graphic images which are used in the application.
  While we do compress these using the Flate compression algorithm, they will increase the PDF file size proportionally.
- The size and number of fonts which are embedded in the statement PDF file will have an impact on the size of the PDF file. The fonts are all in PostScript format. Their sizes can range from 10KB on up to over 100KB. Typically, they average 35,000 bytes.
- In applications where pre-printed paper stock is emulated with an embedded PDF overlay file, the size of the PDF overlay file has an impact. These overlay files should be created with Acrobat Distiller using full compression in the options in order to keep the file size as small as possible.
- Font usage can also have a big impact on the size of the files. We recommend that, when possible, the Acrobat Base 14 fonts be used, since they do not have to be embedded in the PDF file. The fonts include the Time-Roman, Helvetica, Symbol and ZapfDingbats font families. Often, customer Xerox fonts can be mapped to these typefaces and still maintain document look and feel. Because Pro/Meta can use the Xerox font widths to override the PostScript font widths, text positioning and line endings are preserved. Also, fonts which are

not required in the PDF version of the documents can be mapped to the space character of one of the Base 14 fonts.
This eliminates the need for the font. This is often done for bar codes, postal bar codes, control numbers and OMR marks.
The implementer should optimize the size of the statement PDF files before doing any further calculations.

Speed of Conversion

How fast is the metacode statement converted into PDF. For small statements (Less than 10 pages/statement), this conversion is typically less than a second. On larger or more complex statements, this can vary up to 2 or 3 seconds. Conversion speeds will range from 600 to 3,000 pages/minute, depending on statement complexity. Statements with a lot of Page Interleaved Graphic Images are at the lower end of the scale. Speed of the processor and hard disk technology also have an impact on the conversion speed.

There is also a consideration of the amount of embedded resources which need to be added to the PDF file. If a large number of fonts are being embedded, then the time to write these resources into the PDF file must be taken into consideration. This time will vary with the number and size of fonts being embedded, but can add another second or two after the conversion time.

Statement Volume and Frequency

The simple questions of how many statements are being processed, and how often are they created must be answered.

Statement Usage Frequency

Another interesting number to consider is the expected usage frequency of the online statements. Will they be viewed once, or often. When will they be viewed/downloaded - as soon as they are available, during the first week/month, or will usage be spread out in a random nature over a particular period?

Statement Usage Penetration

Another factor is how quickly users will adapt to usage of the online statements. As will any technology, there are early adopters, yet there are many who will be slow to change from their current way of doing things.

The rate of change will vary by user group, and it is important to get an accurate estimate of the rate of change for each application.

Statement Retention Period

How long will the statement files be retained online.

Having gathered the above information, you can calculate the following numbers:

1. The amount of disk space that will be required to store the metacode statement files.
2. The amount of disk space that will be required to store the PDF statement files.

3. The additional disk space required to store the PDF files for the Statement Retention Period.
4. Average end user delay time due to PDF conversion.
5. Batch conversion time for entire statement run.
6. Additional cost of performing batch conversion.
7. Additional cost of PDF file retention.
8. Peak period viewing load.
9. Cost of real-time conversion by putting a value on the resources used, and the users' time.

You can then look at the numbers and make a decision based on the trade-offs of:

Cost of real-time conversion vs. Cost of Batch conversion

This then becomes a simple matter of costs.

Crawford Technologies Inc.
Toll-free number 1–866–679-0864
Telephone 416–410–1467
Fax 416–481–4723
E-mail: support@crawfordtech.com
Website: www.crawfordtech.com