

# Otimizando o Escalonamento de *Jobs* no Processo de Rasterização de Documentos Personalizáveis

Carolina Marques Fonseca, Mateus Raeder,  
Mariana Kolberg, Luiz Gustavo Fernandes

Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS)

Av. Ipiranga, 6681 - Prédio 32 - PPGCC

{carolina.fonseca, mateus.raeder}@acad.pucrs.br, {mariana.kolberg, luiz.fernandes}@pucrs.br

## Introdução

A impressão digital de documentos vem tornando-se cada vez mais eficiente, criando uma nova tendência: a personalização de documentos, que define um *template* a ser instanciado em diferentes situações (por exemplo, um encarte de supermercado que muda suas imagens de acordo com a promoção, mantendo o *layout* do *template*). Com isto, foram criadas linguagens para a descrição de documentos personalizados [PHO 03] e processos para a impressão correta dos mesmos como documentos PDF (*Portable Document Format*), por exemplo. Para personalizar documentos, estas linguagens definem partes estáticas e dinâmicas para os documentos criados.

Atualmente, a maioria das impressoras digitais não é capaz de interpretar as linguagens que apresentam flexibilidade (ou seja, que permitem a definição de áreas estáticas e dinâmicas de um documento). Assim, tornam-se necessários procedimentos adicionais que possibilitem a impressão dos documentos corretamente. Dois destes procedimentos são: a renderização e a rasterização de documentos. Ambos apresentam um alto custo computacional e, se não tratados, representam um gargalo no processo de impressão de documentos em impressoras de alta vazão. Na etapa de renderização, alguns trabalhos envolvendo processamento paralelo já foram realizados para otimizar o desempenho [GFT 06] [FNR 07] [NRK 09]. Já na etapa de rasterização, trabalhos mais recentes vem sendo realizados. O foco do presente estudo é a etapa da rasterização, que é descrita em maiores detalhes a seguir.

A rasterização consiste na conversão do documento em um formato conhecido pelas impressoras, uma vez que elas não são capazes de interpretar as linguagens de alto nível (como PDF e PS) para descrição de documentos. Esta conversão é feita geralmente para o formato *bitmap* (matriz de pontos, com suas correspondentes cores). Através do uso de técnicas relacionadas ao processamento paralelo e distribuído, algumas estratégias já foram aplicadas para aumentar o desempenho desta etapa [NGK 09]. Porém, tais estratégias apresentam alguns problemas, como a impossibilidade de garantir um balanceamento de carga justo para quaisquer sequências de *jobs* contendo documentos personalizados.

Assim, o objetivo deste trabalho é encontrar diferentes tipos de estratégias para obter um bom balanceamento de carga, levando em consideração diferentes características em documentos PDF, como transparência e reusabilidade.

## Visão geral do processo

Para atingir o objetivo deste trabalho, uma ferramenta chamada Roteador Adaptativo de *Jobs* será desenvolvida para buscar características do documento PDF em um arquivo XML (*eXtensible Markup Language*) gerado pelo Profiler [NGK 09]. Após obter as características no arquivo XML, a ferramenta desenvolvida irá gerar um arquivo PS (*PostScript*). Este arquivo servirá como configuração para o RIP (*Raster Image Processor*) - máquina responsável pelo processamento. No arquivo PS estarão especificadas as páginas que cada RIP deve processar. As tarefas dos RIPs devem estar balanceadas da melhor maneira possível, evitando que um RIP com mais trabalho obrigue as outras máquinas a esperá-lo de maneira ociosa. Este balanceamento será baseado em diferentes estratégias, que serão definidas de acordo com as características presentes no documento PDF. A Figura 1 ilustra o processo descrito acima.

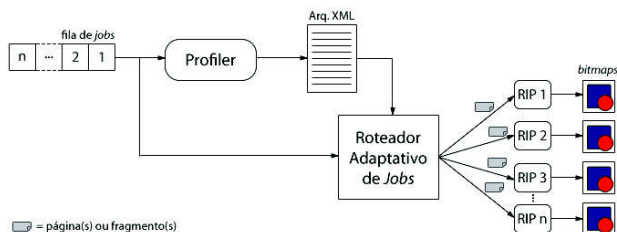


Figura 1: Roteador Adaptativo de *jobs*. Adaptado de [NGK 09].

## Referências

- [GFT 06] Giannetti, F.; Fernandes, L. G.; Timmers, R.; Nunes, T.; Raeder, M.; Castro, M. . High Performance XSL-FO Rendering for Variable Data Printing. In: 21st ACM SAC - Symposium on Applied Computing, 2006, Dijon (França). Proceedings of the 21st ACM Symposium on Applied Computing (ACM SAC). Nova Iorque, NY (EUA): ACM Press, 2006. v.1. p.811-817.
- [FNR 07] Fernandes, L. G.; Nunes, T.; Raeder, M.; Giannetti, F.; Cabeda, A.; Bedin, G. . An Improved Parallel XSL-FO Rendering for Personalized Documents. In: 14th Euro PVM/MPI - European PVM/MPI Users Group Meeting, 2007, Paris (França). Recent Advances in Parallel Virtual Machine and Message Passing Interface (LNCS). Heidelberg : Springer Berlin, 2007. v. 4757. p. 56-63.
- [NGK 09] Nunes, T.; Giannetti, F.; Kolberg, M.; Nemetz, R.; Cabeda, A.; Fernandes, L. G. . Job Profiling in High Performance Printing. In: 9th ACM Symposium on Document Engineering, 2009, Munique. Proceedings of the 9th ACM Symposium on Document Engineering (ACM DOCENG). Nova Iorque, NY (EUA) : ACM Press, 2009. p. 109-118.
- [NRK 09] Nunes, T.; Raeder, M.; Kolberg, M.; Fernandes, L. G.; Cabeda, A.; Giannetti, F. . High Performance Printing: Increasing Personalized Documents Rendering through PPML Jobs Profiling and Scheduling. In: 12th IEEE CSE - International Conference on Computational Science and Engineering, 2009, Vancouver (Canadá). Proceedings of the 12th IEEE International Conference on Computational Science and Engineering (CSE). Los Alamitos, CA (EUA): IEEE Computer Society, 2009. p. 285-291.
- [PHO 03] Purvis, L.; Harrington, S.; O'Sullivan, B.; Freuder, E. C. Creating personalized documents: an optimization approach. In DocEng'03: Proceedings of the 2003 ACM Symposium on Document Engineering, pages 68-77, New York, NY, USA, 2003. ACM.